ER quality control in male fertility

ARTHURS, HAYLEY,MELANIE

ER quality control in male fertility

Hayley Melanie Arthurs

Abstract: Protein disulphide isomerase is a member of the thioredoxin superfamily, resides in the endoplasmic reticulum (ER) and catalyses the formation of disulphide bonds within or between polypeptides, enabling protein quality control. PDILT is a member of the PDI family that specifically resides in the testes and is expressed under developmental control. Recent work has found that, unlike PDI (which possesses the classical CXXC sequence), PDILT has an SXXC domain. This suggests that PDILT does not function fully as a disulphide oxidoreductase but may work as a chaperone. We wished to investigate PDILT and compare it with other PDI proteins, to study the evolution of the PDILT gene and to decipher its protein-protein interactions.

Using sequence analysis of PDILT using the BLAST program, the PDILT protein was compared with other PDI proteins in *homosapiens* and in other species. We then probed further to investigate the expression and interaction of PDILT with other proteins. A transfected cell line model to investigate wt and mutant PDILT was established, and mouse testis tissue lysate was also used to study PDILT interactions. Using immunofluorescence, western blot analysis, and immunoprecipitation techniques, it was established that PDILT forms intermolecular-disulphide complexes and interacts with other chaperones such as CALR3 and CLGN.

Sequence analysis demonstrated conserved amino acids in the b and b’ domains that reflect the specialised roles of the a-a’-b-b’ domains of PDILT. The cysteine residues responsible for inter-molecular disulphide bond formation were identified and found to be important in the interaction of PDILT with CLGN. Lastly, a novel function of the unique tail domain of PDILT was found. The work in this thesis suggests that the tail may be important for the integrity and half-life of the PDILT protein. These studies will help to elucidate the role of PDILT in governing male fertility.

ER quality control in male fertility

Hayley Melanie Arthurs

A Thesis submitted for the degree of Master of Science Department of Biosciences

University of Durham

March 2017

Table of Contents

List of Figures and Tables

Table of Abbreviations

4HB Four Helix Bundle

ADAM3 A Disintegrin and Metalloprotease 3 protein

ATF4 Activating Transcription Factor 4

AZFb Azoospermia Factor b

BiP Binding Immunoglobulin Protein

BSA Bovine Serum Albumin

CALR Calreticulin

CalR3 Calreticulin-3

CLGN Calmegin

CANX Calnexin

CRT Calreticulin

DMEM Dulbecco's Modified Eagle's Medium

DN Diabetic Nephropathy

DTT Dithiothreitol

ECL Enhanced Chemiluminescence

eIF2α eukaryotic translation initiation factor 2α

ER Endoplasmic Reticulum

ERAD ER associated Degradation

ERp46 EndoPDI

Ero1 Endoplasmic Reticulum Oxidoreductin 1

FBS Foetal Bovine Serum

FSH Follicle Stimulating Hormone

GAMPO Goat Anti-Mouse Peroxidase

GPI Glycophosphatidylinositol

GRP94 Glucose-regulated Protein 94

GRP170 Glucose-regulated protein 170

HSP Heat Shock Protein

IP Immunoprecipitation

IZUMO1 Izumo Sperm-Egg Fusion 1

JUNO Folate Receptor 4

LB Lysis Buffer

LH Luteinising Hormone

NEM N-ethylmaleimide

NR Non-reducing

pAB Polyclonal antibody

PBS Phosphate Buffered Saline

PDI Protein Disulphide Isomerase

PDILT Protein Disulphide Isomerase-like Isomerase

PERK PRKR-like ER kinase

PVDF Polyvinylidene Difluoride

R Reducing

RBM RNA binding motif

rER Rough Endoplasmic Reticulum

RNA Ribonucleic Acid

ROS Reactive Oxygen Species

SARPO Swine Anti-Rabbit Peroxidase

SDS Sodium Dodecyl Sulphate

sER Smooth Endoplasmic Reticulum

STRING Search Tool for the Retrieval of Interacting Genes/Proteins

TBS Tris buffered saline

TBS-T Tris Buffered Saline with Tween

TEMED N, N, N', N'-Tetramethylethylenediamine

TXNDC11 Thioredoxin Domain Containing 11

UPR Unfolded Protein Response

UTJ Uterotubal junction

WT Wild Type

XOMAT X-ray Developer Machine

Declaration

I declare that the contents and all data presented in this thesis, is the result of my own work. No part of the material offered has previously been submitted for a higher degree. The body of work has been achieved under the supervision of Dr. Adam M. Benham.

Hayley Melanie Arthurs

March 2017

Statement of Copyright

Acknowledgments

I would like to take this opportunity to thank Dr Adam Benham for giving me the opportunity to work on this project and giving guidance throughout. I would also like to thank all the members of Lab 8 for assisting me and giving me helpful advice.

Finally, I would like to thank my family for funding me throughout this project and for their support.

1. Introduction

1.1 Structure and function of the testis

The testes are major organs which play a significant role in male fertility. The primary function of the testis is spermatogenesis and to produce androgens, primarily testosterone. Testosterone is a naturally occurring steroid hormone which serves an important role in the endocrine system in the testis. This function is essential for the development of the male reproductive system. Within the structure of the testis, there are fine coiled tubes called *seminiferous tubules*. The seminiferous tubules are approximately 50 cm in length and 200 µm in diameter (Green, O, Stout, G. W, Taylor, D. J., 1985). It is within the location of these seminiferous tubules where spermatogenesis occurs. These tubules are also embedded in connective tissue containing Leydig cells which produce testosterone when stimulated by the luteinising hormone (LH). Spermatogenesis occurs through several stages: proliferation, growth, and maturation. The proliferation stage occurs through many divisions of spermatogenic cells by mitosis (figure 1). The spermatogenic cells are contained within a layer of germinal cells also known as seminiferous epithelium (Peckham, M., *et al.,* 2003).

Studies of the seminiferous epithelium have helped researchers to understand the dynamics of renewal and proliferation of spermatogenic cells. For instance, Amman (2008) revisits what is known about the 'cycle of the seminiferous epithelium', such as the renewal and proliferation of spermatogonia and the ways in which the deficiencies in data can be rectified. It is also important to determine when

spermatogonia become committed to proliferate or how many mitotic divisions occur thereafter (Amman, R. P., 2008). Understanding how the seminiferous epithelium is maintained would enable evaluation of the quantity and the quality of sperm ejaculated, which demonstrates another perspective of investigating male fertility.

After the proliferation stage, spermatogenic cells undergo a period of growth, in preparation for the last stage maturation. The final stage involves the two meiotic divisions to form gametes with half the genetic material (23 N) and the differentiation of these haploid cells to form spermatids.

Along with spermatogenic cells the seminiferous epithelium contains sertoli cells. Sertoli cells are the epithelial supporting cells of the seminiferous tubules. They play an important role in regulating the process of spermatogenesis through mechanical support, protection, and nourishment during the maturation stage. They do this through the transduction of signals from testosterone and Follicle Stimulating Hormone (FSH) for the successful production of the sperm. They also allow the exchange of important factors such as nutrients and growth factors which are required to fuel germ cell metabolism and therefore the growth and maturation of the spermatozoa (Walker, W. H, Cheng, J., 2005).

Testosterone ($C_{19}H_{28}O_2$) is a potent androgenic steroid which is made primarily in the testes found in the male reproductive system. Leydig cells produce testosterone when stimulated by the luteinising hormone (LH) from the pituitary gland (Walker, W. H, Cheng, J., 2005). Testosterone provides feedback control of the LH and FSH hormones secreted from the pituitary gland. This feedback control mechanism is necessary for the development of the spermatogonia. Testosterone is also necessary for the development of male sex characteristics such as muscle growth, facial hair,

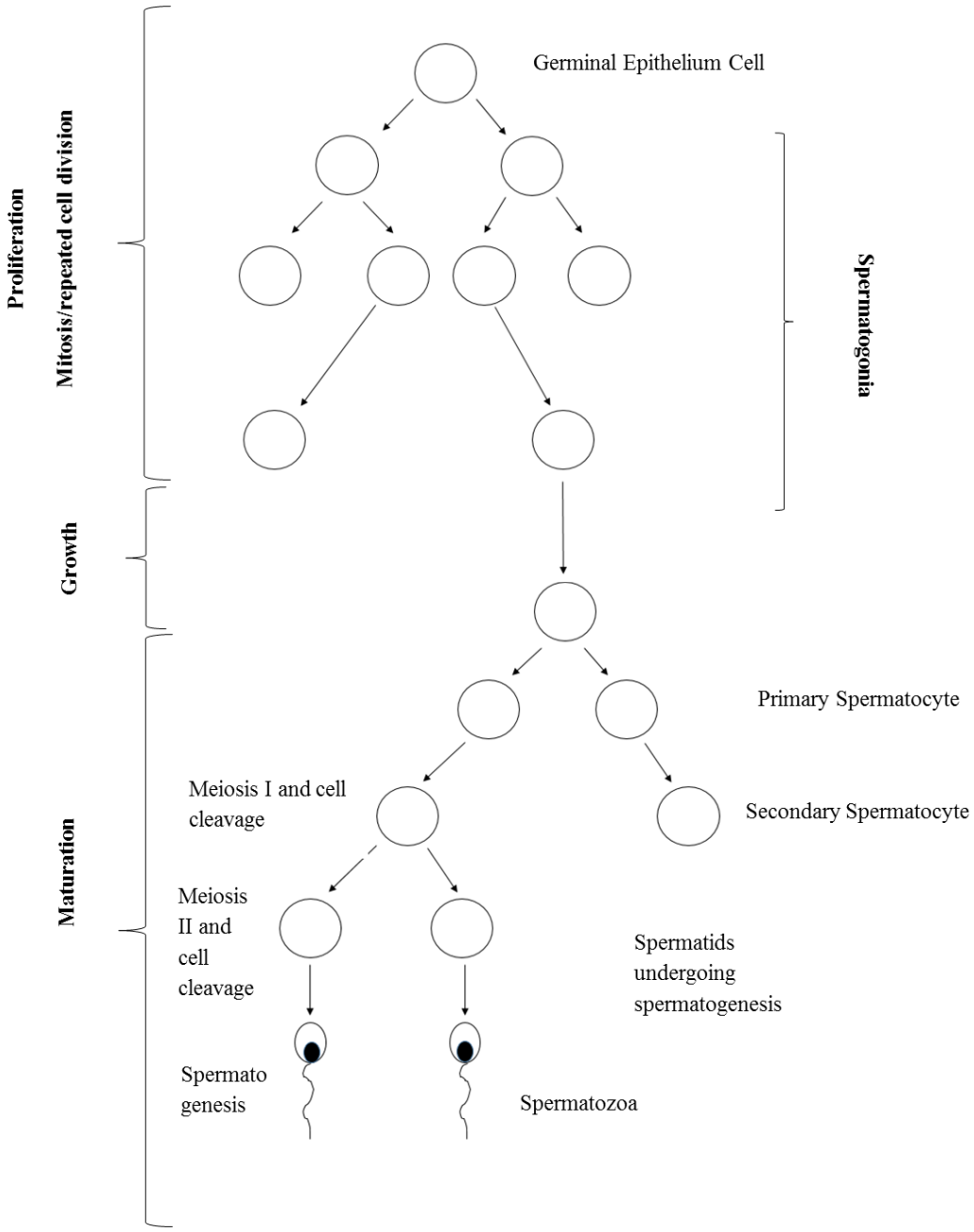


Figure 1: Spermatogenesis – The diagram shows the three stages of spermatogenesis: proliferation, growth, and maturation from a single germinal epithelium cell. These stages occur in the seminiferous tubules.

and a deep voice. Ageing is also associated with decreased levels of testosterone (Rajfer, J., 2003), which can lead to erectile dysfunction. Follicle-stimulating hormone (FSH) is a glycoprotein polypeptide hormone which is required for the determination of Sertoli cell number, and for induction and maintenance of normal sperm production (Simoni, M., *et al.,* 1999). Therefore, FSH plays an important role in spermatogenesis.

Spermatozoa are highly specialized cells with specific metabolic pathways compartmentalized in different regions. The metabolic control of ATP production is key to the function of sperm motility and maintaining protein post-translational modifications such as phosphorylation. During spermatogenesis, different metabolic pathways support specific processes in differentiation of germ cells. For example, during epididymal maturation the glycolysis pathway is important (Miki, K., 2007). However, while the acrosome reaction takes place, oxidative phosphorylation is a major driver for ATP production. Unlike other many cells, spermatozoa are exceptionally differentiated and show large variations in their genetic, cellular, and chromatin structures, which allow them to control fertility/infertility, embryonic development and heredity.

The head structure contains a nucleus, with the haploid number of chromosomes. It also contains the acrosome structure over the head of the nucleus, which is a membranous organelle. It is a highly-conserved organelle, which contains many hydrolytic enzymes which enables the penetration of the oocyte prior to fertilization (Berruti, G, Paiardi, C., 2011). One of the main type of enzymes that are important for fertilization are hyaluronidases (EC 3.2.1.35). They are released by the acrosome after they have reached the oocyte, digesting hyaluronan in the corona radiata, thus enabling conception. Another type of enzyme important for fertilization are

Acrosins. They are digestive enzymes which act as proteases. They are released from the acrosome and aid the penetration of the zona pellucida, enabling the acrosome reaction.

At the neck of the spermatozoa are two centrioles, and the microtubules of one centriole extends to the rest of the spermatozoa to the flagellum (Green, O, Stout, G. W, Taylor, D. J., 1985). The flagellum is an important feature as it enables the spermatozoa to propel itself through the female tract, cluster around the oocyte and orientate the spermatozoa so that it is ready to penetrate the oocyte for fertilization. Eukaryotic flagella are highly specialized for locomotion and consist of cytoskeletal structures called axonemes. The axonemal bend movement, which enables locomotion, occurs through the active sliding of the two microtubules by the molecular motor dynein (Inaba, K., 2011). The axoneme is also surrounded by outer dense fibres as well as mitochondria to power the active sliding of the dynein.

The molecular mechanism in which the sperm binds to the egg has been somewhat of an enigma (Klinovska, K, Sebkova, N, Dvorakova-Hortova, 2014). A few molecules are now known to play a role in this mechanism, including the key sperm protein IZUMO 1. It was found that the knockout of this molecule led to infertility in males only (Inoue, N., *et al.,* 2005). This molecule appears to be present on the sperm only, and not the oocyte. In contrast, another molecule that has been found to take part in the binding of the sperm to the egg is a folate receptor 4 called ‘Juno’. It has been found to reside on the oolema and binds to IZUMO 1 specifically (Bianchi, E., *et al*., 2014) (chapter 1.13 for more details).

The extracellular receptor pair (IZUMO1 and Juno) are the only known extracellular receptor pair that allow the gametes to bind successfully. The binding of this receptor

pair has been found to be necessary for the adhesion event between the sperm and egg for fertilization (Bianchi, E., *et al*., 2014). However, this receptor pair isn't necessary for the fusion event of the sperm with the egg and more research is needed to identify the fusogenic proteins and their specific roles in the fusion of the two gametes. Lastly, tetraspanin CD9, is also present on the egg and is indispensable in fertilization. However, the molecular role of CD9 is still unknown (Jankovicova, J., *et al.*, 2014).

A recent paper carried out a systematic review, meta-analysis, and population-based study, it was calculated that at least 30 million men are infertile worldwide, with the highest rates in Africa and Eastern Europe (Agarwal, A, *et al.*, 2015). There are several known causes of male infertility and these include: oligospermia, asthenospermia, teratozoospermia and azoospermia. These causes were found in 20–25% of cases (Egozcue, *et al*., 2000; Hargreave, 2000). Oligospermia, for example, is a condition where there is a low concentration of sperm in the semen. There are many potential factors that can lead to oligospermia. The factors include lifestyle habits such as smoking, alcohol and drugs which can lead to the inability to produce hormones to trigger spermatogenesis in the testes. There are also testicular factors, such as genetic defects on the Y chromosome, age and trauma, which can lead to low production of sperm. Lastly, there are factors that can affect the male genital system after sperm production and impede the ejaculation of sperm, which include vas deferens obstruction, infections such as prostatitis (inflammation of the prostate gland), or lack of vas deferens (which is found in people with cystic fibrosis).

Another known cause of male infertility is asthenospermia, which is essentially reduced sperm motility. One in five thousand men have complete asthenospermia, where all sperm present are immotile, and this can be found in patients with

metabolic deficiencies or abnormalities in the sperm flagellum (Ortega, C., *et al.*, 2011). Another known cause of male infertility is teratozoospermia. Teratozoospermia, is a condition where the sperm displays abnormal morphology. The factors that lead to this condition are mostly unknown, but it is found to be a common characteristic in certain diseases such as Hodgkin's disease, coeliac disease, and Crohn's disease. Lifestyle factors such as smoking can also cause this condition (Martin-Du Pan, R., *et al.*, 1997).

Lastly, azoospermia is another cause of male infertility when the sperm count is equal to or less than 15 million/ml in the semen. The causes of this can be due to reproductive tract obstruction (obstructive azoospermia) or inadequate production of spermatozoa (non-obstructive azoospermia). Factors that have been found to be the cause of obstructive azoospermia are congenital (absence of vas deferens or idiopathic epididymal obstruction), through infections or by a vasectomy (Schlegel, P. N., 2004).

More research into unexplained male fertility is needed, and many factors such as cultural barriers and social stigma prevent progress in research. Our knowledge and understanding of spermatogenesis is increasing; however, it is a complex process and it is estimated that about 2000 genes regulate spermatogenesis (Hargreave, 2000). Most of these genes are present on the autosomes (not on a sex chromosome). However, there are approximately 30 genes present on the Y-chromosome. One family of genes found on the Y chromosome is the RNA binding motif family (RBM). There are up to 50 genes that are members of the RBM family. RBM is a nuclear protein, which likely metabolises RNA, and is specifically expressed in the male germ line in humans and mice. It appears that RBM regulates splicing events which are essential for spermatogenesis. RBM has been found to be inactivated

during deletions of the azoospermia factor b (AZFb) regions. The AZF region is located on the human male Y chromosome, deletions in this region have been associated with the inability to produce sperm.

Another gene that is found on the Y chromosome and regulates spermatogenesis is the DAZ gene. It is a member of a 6-10 gene family, and like RBM it is likely involved the metabolism of RNA. However, unlike RBM the DAZ protein is cytoplasmic. Both RBM and DAZ proteins are related. DAZ, however, is important for gametogenesis in both sexes. They may be involved in regulating translational repression during spermatogenesis, but more research is needed to explore this area. There is also little known concerning the fusion proteins that are responsible for the sperm accessing the egg. A greater knowledge of the processes behind male fertility has potential to enable new therapies for infertile men, and their partners, who want children and to discover new methods of contraception.

1.2 The endoplasmic reticulum

The endoplasmic reticulum is a membranous organelle which mainly consists of continuous membranes, a network of tubules and sacs called cisternae (Cooper, G. M., 2000). New super-resolution imaging technologies have recently been used to elucidate the morphology and dynamics of the peripheral ER (Nixon-Abell, J., *et al.*, 2016). They found that the ER consists of tubules of varying densities, including structures called ‘ER matrices’. These structures were originally misidentified as sheets because of the dense clustering of tubular junctions and ER motion. The presence of ER matrices is dynamic and therefore may enable the ER to be more flexible and change its conformation to adapt to cellular needs.

There are two types of ER membranes: the rough ER (RER), and the smooth ER (SER). The rough ER is called 'rough' because it has ribosomes attached to the membrane, serving a function in protein translation and targeting. The smooth ER does not bind ribosomes and therefore serves a different role, which is to synthesise lipids, phospholipids, and steroids. The testis contains a large amount of smooth ER because it produces steroids such as testosterone.

One of the most important functions of the RER in eukaryotic mammalian cells is maintaining quality control of proteins. The RER is a specialised organelle where proteins are translocated to and start to fold co-translationally. Folding is then completed post-translationally. However, for some proteins, the process of protein folding, and maturation is slow and needs post-translational modifications in addition to signal sequence cleavage, such as disulphide bonds (section 1.5), proline isomerisation and N-linked glycosylation (Ellgaard, L, Helenius, A., 2003).

The RER has a specialised structure and environment, which enables it to function in protein quality control. One of the ways in which the ER is adapted for protein folding is that it is more oxidising than the cystol. A relatively oxidising environment is necessary for protein folding through disulphide bond post-translational modifications. Another important feature of the ER is it contains specialized chaperones and enzymes. There are many different types of chaperones present in the ER that help to fold proteins. There are general chaperones, such as GRP78/BiP, GRP94 and GRP170, which are members of the heat-shock protein family. There are also lectin-like chaperones, such as Calnexin and Calreticulin, peptidyl-prolyl isomerases such as PPI and lastly redox active chaperones such as PDI and ERp57. During folding, some chaperones can recognise specific features which are exposed during folding. For example, HSP70s such as GRP78 recognise

hydrophobic surfaces on proteins during folding. The lectin-like chaperones such as Calnexin and Calreticulin recognise immature glycans and indirectly regulate the retention of incompletely folded secretory glycoproteins in the ER or their movement onto the Golgi (Ellgaard, L, Helenius A., 2003, Ellgaard, L, Molinari, M, & Helenius, A., 1999). By forming transient complexes with nascent proteins, specific folding factors, chaperones, and enzymes ensure that only correctly folded proteins are released.

1.3 ER Chaperones

As mentioned above, ER molecular chaperones aid the folding of soluble and integral membrane secreted proteins. They prevent protein aggregation and are foldases that directly promote folding, quality control, and export from the ER. There are three groups of molecular chaperones including: heat shock proteins, lectin-like chaperones, and thiol oxidoreductases of the PDI family, which will be discussed later (Nishikawa, S, Brodsky, J. L, Nakatsukasa, K., 2005).

Heat-shock proteins (HSP) are highly conserved molecular chaperones and play roles not only in protein folding but also in the transport/sorting of proteins, cell-cycle control and signalling and protection of cells against stress/apoptosis (Li, Z, Srivastava, P., 2004). They are often classified based on their molecular weight. For example, GRP78/BiP is an ~70 kDa HSP that is involved in the translocation, folding, and assembly of secretory and transmembrane proteins (Kohno, K., *et al.*, 1992). Glucose-regulated protein 94 (GRP94) is a HSP90-like protein which is found in the lumen of the ER which chaperones secreted and membrane proteins (Marzec, M., *et al.*, 2012). Lastly, Glucose-regulated protein 170 (GRP170) is an HSP70-like protein that is involved in assisting protein folding, targeting misfolded

proteins for degradation and in regulating the transducers of the Unfolded Protein Response (Behnke, J, Hendershot, L. M., 2013).

Calnexin and Calreticulin proteins act as an ER lectin chaperone complex system, which helps direct the maturation of glycoproteins in the ER. These lectin chaperones recognise $Glc_1Man_9GlcNAc_2$ specifically on the glycoprotein intermediate seen during early oligosaccharide processing (Williams, D. B., 2005). The ER lectin chaperones also possess binding sites for ATP, calcium ions, non-native polypeptides and ERp57, a PDI homolog that catalyses disulphide bond formation, reduction, and isomerization.

1.4 The Unfolded Protein Response

The correct folding of proteins is essential, as aberrant proteins are extremely harmful to cells. If aberrant proteins accumulate in the cystol, this triggers the heat-shock response, which leads to the elevated synthesis of HSP70 and other cytosolic chaperones. However, if aberrant proteins accumulate in the ER, this can result in ER stress (Sitia, R, Brakkman, I., 2003). Upon ER stress, unfolded and misfolded proteins bind to the ER resident heat shock protein HSP70 (BiP) and activate the unfolded protein response (UPR). There are three signalling pathways that originate from the UPR (figure 2).

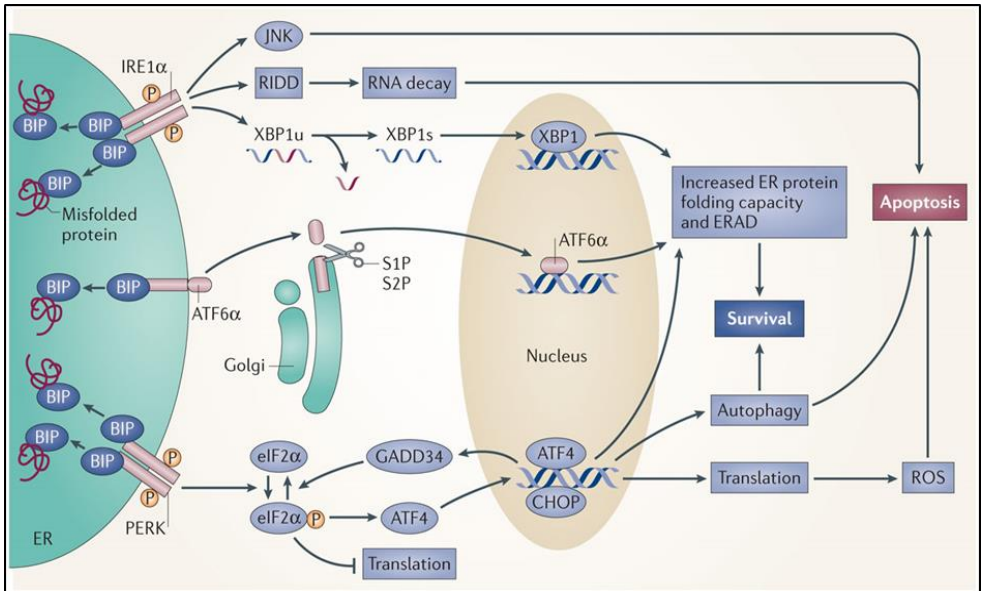


Figure 2: UPR signalling pathways (taken from Wang, M, Kaufman, R. J., 2014). This diagram shows the three signalling pathways (Ire1α, ATF6α and PERK) that become activated because of ER stress and result in the UPR. When ER stress occurs the misfolded protein binds to BIP immunoglobulin heavy-chain binding protein (BIP), and becomes sequestered. It then can trigger any of the three signalling pathways. For instance, two BIP proteins and two phosphate molecules could bind to the inositol-requiring protein 1α (IRE1α) and the first signalling pathway would take place. Or a single BIP protein can bind to ATF6α which would lead to another signalling pathway. Finally, two BIP proteins and two phosphate molecules can bind to PPKR-like ER kinase (PERK) which would lead to the third signalling pathway. As a result, increased protein folding, transport, and ER-associated protein degradation (ERAD), as well as attenuating protein synthesis. However, if this doesn't resolve the issue it ultimately leads to apoptosis of cells.

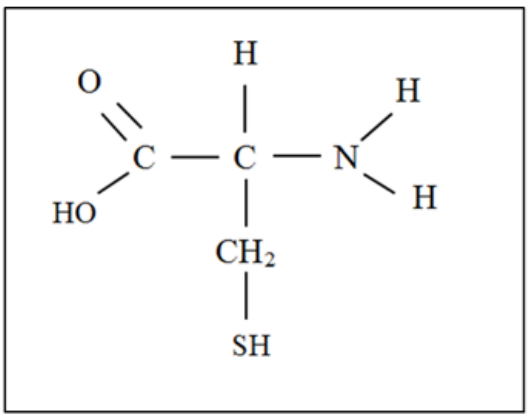


Figure 3: Cysteine residue. As with any amino acids this residue contains the standard carboxylic acid (on the left), a hydrogen atom (on top) and an amino group (on the right). What makes the cysteine residue unique is its side chain contains a thiol group (SH), which is attached to a methyl group.

In the first signalling pathway, BiP binds to PRKR-like ER kinase (PERK), which catalyses the transfer of phosphate onto the eukaryotic translation initiation factor 2α (eIF2α). This in turn inhibits translation and therefore produces cell cycle arrest in the G1 phase. Activating transcription factor 4 (ATF4) is then activated which binds to genes that upregulate UPR. Overall, this leads to the production of ROS and potentially autophagy, and ultimately leading to apoptosis. It also causes increased ER folding capacity and ER associated Degradation (ERAD), which leads to survival. Lastly, ATF4 causes a feedback loop through activation of GADD34 which dephosphorylates eIF2α (Wang, M, Kaufman, R. J., 2014).

In the second signalling pathway, BiP is released from ATF6α, which results in its transport into the Golgi, where it is cleaved by proteases, releasing the transcription factor domain. The cytosolic portion of ATF6 then translocates to the nucleus and binds to stress element promoters upstream of genes that upregulate UPR, therefore increasing the ER protein folding capacity and ERAD.

In the final signal pathway BiP is released from IRE1α upon binding of unfolded proteins; Ire1a activates JNK and RIDD, which leads to apoptosis if the UPR is unresolved. Ire1 also activates the transcription factor XBP1 by alternative processing of its mRNA. UPR stress genes are in turn upregulated by binding to stress element promoters in the nucleus and increased ER protein folding capacity and ERAD occurs (Sitia, R, Brakkman, I., 2003). However, excessive and prolonged ER stress triggers cell suicide, usually in the form of apoptosis, representing a last resort of multicellular organisms to dispense of dysfunctional cells by apoptosis (Xu, *et al.*, 2005).

1.5 Post-translational modification - disulphide bonds

For proteins to be translocated into the ER, most must be translocated via the ER-translocon. However, fully folded proteins cannot fit through the translocon. Therefore, post-translational folding occurs in the ER. One of the key post-translational modifications in the ER is the provision of disulphide bonds. It has been estimated that disulphide bond formation is involved in the biogenesis of approximately one third of human proteins. Their usual function is to stabilize protein native structures, although they may also play other roles, including redox regulation of enzymatic activity (Hatahet, F, Ruddock, L. W., 2009).

Disulphide bond formation can only occur in proteins with cysteine residues that are readily available and are in close proximity, as on average the length of a disulphide bond is approximately 2Å between two cysteine residues. The thiol group (-SH) on the side chain of the cysteine residue, is the location where disulphide bond formation takes place (figure 3).

The thiol group has a -2-oxidation state, however, when two thiol groups bond they are oxidized to form a disulphide (figure 4). The sulphur atoms found in the products of the reaction have a -1-oxidation state each. Two electrons and two protons are released. For example, in a simple reaction, molecular oxygen would act as an oxidant, as it would uptake the two electrons and protons, thus forming hydrogen peroxide (Hatahet, F, Ruddock, L. W., 2009).

Disulphide bond formation is one of the key rate-limiting steps in protein folding of many proteins. This suggests that in many proteins, catalysis and regulation of disulphide-bond formation is crucial for folding. This post-translational modification is commonly found in secreted proteins, which enter the ER with reduced cysteines and leave it with oxidized cysteines (figure 5).

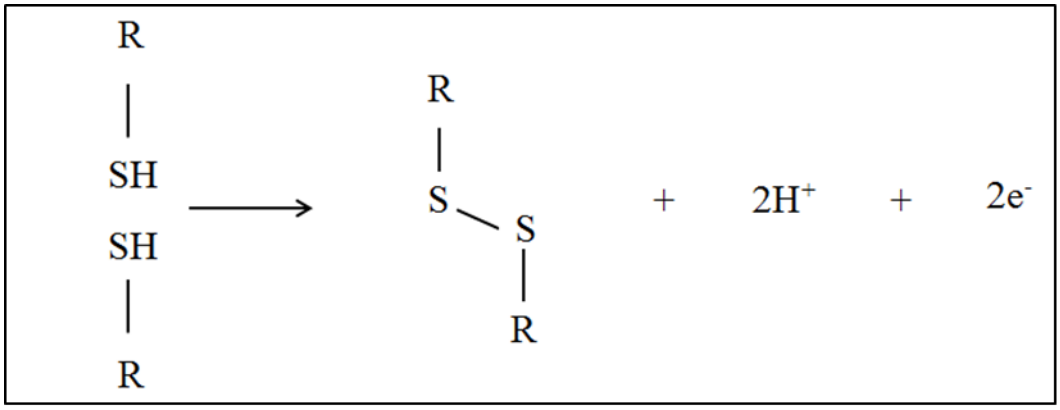


Figure 4: Formation of a disulphide bond. The reaction is an oxidation reaction between two thiol groups, which get oxidised to form a disulphide bond with the loss of two electrons and protons.

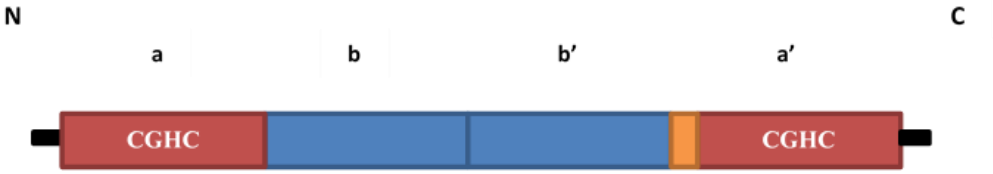


Figure 5: Structural representation of PDI. This diagram represents the domains. The 'a-type domains' are represented in red and they are the active sites. The cysteine residues in the CXXC motifs, labelled in the 'a-type domains', form disulphide bonds. The b-type domains are represented in blue and the linker between the b' and a' is represented in orange.

1.6 Protein Disulphide Isomerase (PDI)

PDI is a member of the thioredoxin superfamily and there are at least 21 PDI family members present in mammalian cells, most of which can act as disulphide exchange proteins (Benham, 2012). PDI contains four thioredoxin domains including two a-type domains, which contain the active site with CGHC residues (Lith, M. V., *et al*., 2007), (figure 5). The cysteine residues found in the CXXC motif of the active site shuttle between the reduced (dithiol) and oxidized (disulphide) form (figure 4). Whether the active site cysteines exist in the reduced or oxidized state depends on the stability of the disulphide and the ability of an ER oxidase to catalyse its oxidation. There are also two b-type domains, which lack the active site, and therefore lack redox activity, but have either a structural function or are involved in substrate recognition and binding. These domains have an a-b-b'-a' organisation. PDI resides in the ER and catalyses the formation of disulphide bonds within or between polypeptides entering the secretory pathway in the ER, enabling efficient protein folding (figure 6). The outcome of these reactions depends on the relative stability of the disulphide in the enzyme and substrate. To catalyse the formation of disulphide bonds, PDI takes part in a series of disulphide exchange reactions. However, PDI needs to be oxidised first by specific oxidases (Bulleid, 2012). An example of an oxidase that oxidises PDI is the ERO1 protein. There are two Ero genes present in mammals (ERO1a and ERO1b) and one that is present in *S. cerevisiae* (encoding Ero1p). A series of oxidation/disulphide exchange reactions between ERO1 and PDI (from dithiol to disulphide), allows the oxidation of the substrate (figure 7) and therefore the formation of disulphide bonds in the client polypeptide (Onda, 2013).

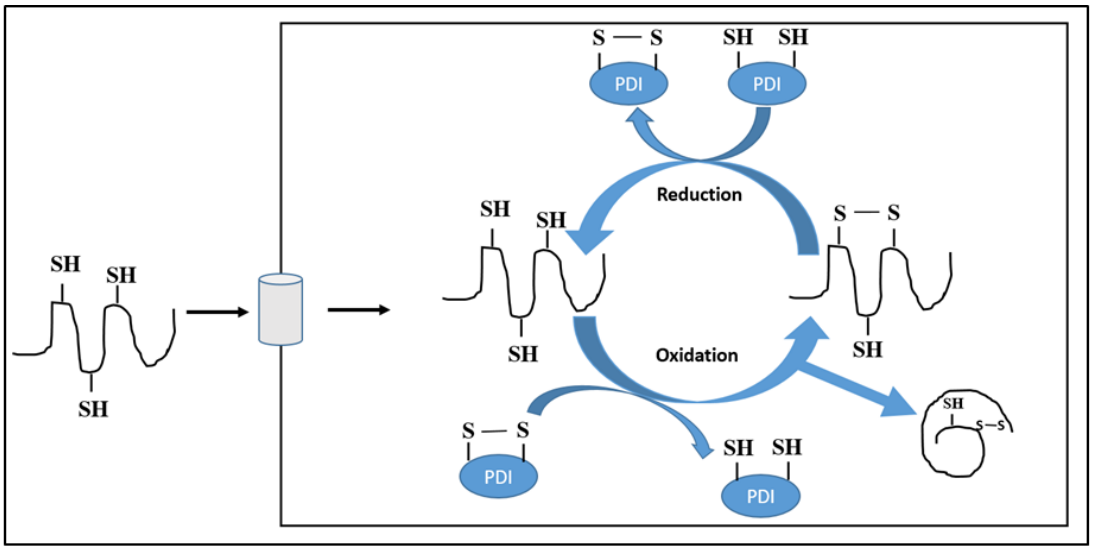


Figure 6: Folding of the polypeptide within the ER. Polypeptide chains are co-translationally translocated across the ER membrane by the sec61 ER translocator. The formation of disulphide bonds is then catalysed by PDI between cysteine residues. The cysteine residues shuttle from oxidised to reduced form, causing the oxidation of the disulphide bonds of the polypeptide chain (Bulleid, 2012).

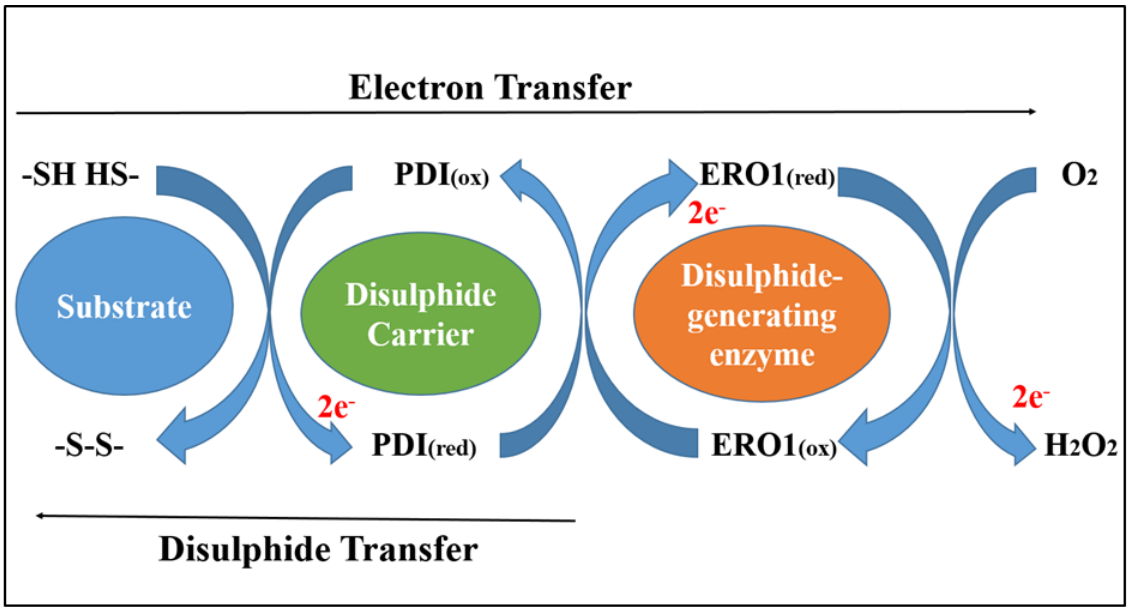


Figure 7: Disulphide transfer between oxidoreductases. ERO1 is oxidised by oxygen releasing two electrons and protons to form hydrogen peroxide. ERO1 is then able to accept two electrons from PDI, which oxidises PDI allowing it to in turn oxidise the substrate through accepting two electrons. The oxidation of the substrate enables the formation of disulphide bonds (Onda Y, 2013).

1.7 PDI homologs and their function

As mentioned in the previous section, there are at least 21 PDI homologs in mammalian cells with a variety of functions (table 1). There are many PDI members included in this table that have unknown physiological functions and more research is needed to elucidate their role.

The PDI family is diverse; some members are ubiquitously expressed in many different cells, such as PDI itself, ERp57 and P5. ERp57 has a great number of functions within and outside the ER (Hatahet, R, Ruddock, L. W., 2009). ERp57 interacts with Calnexin and Calreticulin chaperones and promotes the oxidative folding of glycoproteins (Frickel, E. M., *et al.*, 2004). It also plays an important role in the assembly of MHC class I molecules by facilitating peptide loading and working with tapasin. P5 is normally ER localised because, like all PDIs, it has a KDEL sequence and can interact with other ER chaperones such as PDI, ERp77 and BiP. The presence of P5 has also been found extracellularly, outside the ER and plays a role on the cell surface modulating the oxidation state of integrins involved in blood clotting. P5 has also been shown to inhibit the cytosolic face of Sec61 ER translocator protein (figure 6), (Stockton, J, Merkert, M. C, Kellaris, K. V., 2003).

There are other members of the PDI family that are expressed exclusively in specific tissues. PDIp is expressed in the acinar cells of the pancreas, and functions as a protein-folding catalyst for secretory digestive enzymes. PDILT is not only expressed specifically in the testes but is also under developmental control because it only becomes expressed after puberty in post mitotic spermatids (Lith, M., *et al.*, 2007). Lastly, EndoPDI (ERp46) is expressed in endothelial cells and is upregulated during hypoxia (Sullivan, D. C., *et al.*, 2003). The suggested role of ERp46 may be an important component in the ‘glucose toxicity’ involved in insulin production, at

the post-translational level (Alberti, A., *et al*., 2009). 'Glucose toxicity' is a term used for a decrease in insulin secretion and an increase in insulin resistance because of chronic hyperglycaemia that is found in type 2 diabetes.

Lastly, there is a group of PDI-family members that are transmembrane spanning. There are at least 5 of these. Of the 5, TMX1, TMX2 and TMX3 have been shown to have a putative type I transmembrane protein ER localization signal (Hatahet, R, Ruddock, L. W., 2009). However, more research is needed, as the physiological function of TMX4 and TMX5 hasn't been elucidated yet.

Table 1 (part 1): Human PDI family The table shows the size, structure, active site sequence and function of the PDI family (Adapted from Hatahet, R., Ruddock, L. W, 2009 and Appenzeller-Herzog, C., Ellgaard, L., 2008).

Name	Length	Domains	Number of a type domains	Active site sequence	Functions
PDI	508	a-b-b'-a	2	CGHC	Catalyst of thiol-disulphide exchange reactions
PDIp	525	a-b-b'-a	2	CGHC, CTHC	Expressed specifically in the ancinar cells of the pancreas
PDIr	519	b-a°-a-a'	3	CSMC, CGHC, CPHC	Unknown, androgen (270) and Hox (171)-regulated gene
PDILT	584	a-b-b'-a	2	SKQS, SKKC	Expressed specifically in the testes after puberty
P5	440	a°-a-b	2	CGHC	Catalyst of thiol-disulphide exchange reactions and upregulated by XBP-1 for UPR
Hag 2	175	a	1	ECPHS	Structural homology with ERp18, function unknown but are present in breast tumours
Hag 3	165	a	1	DCQYS	Structural homology with ERp18, function unknown but are present in breast tumours
ERdj5	793	J-a''-b-a°-a-a'	4	CSHC, CPPC, CHPC, CGPC	ER Reductase that Catalyses the Removal of Non-Native Disulphides and Correct Folding of the LDL Receptor.
ERp18	172	a	1	CGHC	Unknown
ERp27	273	b-b'	0	-	Unknown – but binds to polypeptide with ERp57

Table 1 (part 2): Human PDI family The table shows the size, structure, active site sequence and function of the PDI family (Adapted from Hatahet, R., Ruddock, L. W, 2009 and Appenzeller-Herzog, C., Ellgaard, L., 2008).

Name	Length	Domains	Number of a type domains	Active site sequence	Functions
ERp27	273	b-b’	0	-	Unknown – but binds to polypeptide with ERp57
ERp28/29	261	b-D	0	-	Expressed during ER stress, control of oxidative protein folding
ERp44	406	a-b-b’	1	CRFS	Expressed during ER stress, control of oxidative protein folding
ERp46 /EndoPDI	432	a°-a-a’	3	CGHC	Expressed specifically in endothelial cells. Protect cells from hypoxia
ERp57	505	a-b-b’-a’	2	CGHC	Oxidative folding of glycoproteins
ERp72	645	a°-a-b-b’-a’	3	CGHC	Catalyst of thiol-disulphide exchange reactions
ERp90[a]	825	a-b-b’	-	-	ER FAD interactor
TMX*	280	a	1	CPAC	Oxidoreductase activity and this may help relieve ER stress
TMX2*	296	a	1	SNDC	Unknown
TMX3*	454	a-b-b’	1	CGHC	Protein dithiol-disulphide oxidant, and the recombinant luminal portion of the protein has reduction potential similar to that of PDI
TMX4*	349	a	1	CPSC	Unknown
TMX5*	-	a	1	CRF5	Unknown

*Transmembrane PDI- family members, [a] (Riemer, J. *et al.* 2011)

1.8 Protein Disulphide Isomerase-Like Protein of the Testis (PDILT)

PDILT is a member of the PDI family. It is unique because it is not only expressed specifically in the testes, but it is also the only known PDI homolog that is under developmental control. It is expressed in post-meiotic male germ cells and is induced during puberty (Lith, M., *et al*., 2007).

Human PDILT has two cysteine residues: one in the catalytic site of the second catalytic domain (C420) and one in the first catalytic domain (C135) but away from the active site (figure 8). The free cysteine (C135) is not found in other PDI family members and may be positioned there to enable specific substrate interaction. C420, found within the a' SXXC motif, also engages in disulphide bond dependent interactions with unknown proteins (Lith, V. M., 2005). However, PDILT lacks the classical CGHC sequence, so cannot be a fully functional oxidoreductase, although radiolabelled peptide and hormone binding studies have shown that PDILT has the capacity to work as a chaperone (Lith, M., *et al.*, 2007).

As a PDI homolog, PDILT shares 27% of sequence identity with PDIA1, and the a' domain is the most strongly conserved between these proteins with 48% identity. However, as mentioned previously, PDILT is an unusual PDI with unique features that set it apart from other PDIs. One of the most striking differences, , lies within the catalytic a and a' domains. The 'so-called' active site possesses SKQS and SKKC motifs that have diverged from the classical CXXC motif. The absence of catalytic cysteines explains why PDILT may primarily function as a chaperone (Bastos-Aristizabal, *et al.*, 2014).

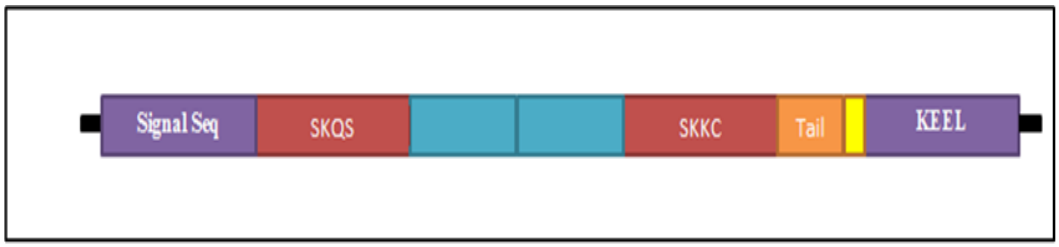


Figure 8: Domain organisation of PDILT. The red domains are the catalytic 'a domains' which have a divergent active site and the blue domains represent the 'b domains', which have the client binding residues. There is a unique hydrophobic tail sequence (orange) and the purple domain has a KEEL sequence which enables ER retention.

1.9 PDILT localization and tissue expression

PDILT is an ER resident glycoprotein. To demonstrate this, HeLa cells were transfected with cDNA encoding the PDILT protein with a Myc tag upstream of the putative ER retention motif (PDILT-Myc) and were analysed using immunofluorescence confocal microscopy (Lith, M. V., *et al.*, 2005). PDILT co-localised with Ero1α, a known ER resident protein. These findings were further verified by determining the glycosylation status of PDILT-Myc (as PDILT has 9 N-glycosylation sites and N-glycosylation occurs in the ER), by treating the cell lysates with Endoglycosidase H (which digests immature N-glycans but not modified ones found in the Golgi) and analysing the molecular weight of PDILT-Myc by SDS-PAGE and Western blotting. Increased mobility of PDILT-myc was observed, showing that PDILT-Myc was N-glycosylated (Lith, M. V., *et al.*, 2005).

PDILT is expressed in post-meiotic male germ cells (Lith, M., *et al.*, 2007). An anti-serum raised by the group against PDILT specifically recognised both transfected PDILT in HeLa cells and endogenous PDILT in mouse testis. The antiserum didn't detect PDILT in non-testicular tissues. Furthermore, immunohistochemistry was conducted on adult rat testis sections; testis sections were stained for PDILT and counterstained with haematoxylin to determine the cellular site of PDILT expression. It was found that PDILT was present in the most mature germ cells located around the seminiferous tubule lumen. Immunopositive staining was first detected in haploid round spermatids found at stage VII. The onset of PDILT expression commenced upon completion of meiotic prophase. Thereafter, PDILT was present in the cell body of all subsequent round and elongating spermatids.

PDILT expression is under developmental control. RNA isolated from rat testes at different ages was used in RT-PCR to determine when PDILT is expressed (Lith, M.

V., *et al*., 2007). It was found that PDILT mRNA was detected in testis only from day 30 and reached adult levels from day 45 onwards. This was confirmed at the protein level by western blotting. Immunohistochemistry of testes from rats at different ages was also performed to show that expression of PDILT corresponded with the formation of round spermatids during the first spermatogenic wave that takes place during puberty. On day 31, some of the seminiferous tubules from rat testis demonstrated expression of PDILT. The expression of PDILT in the seminiferous tubules increased dramatically up to day 45, where all seminiferous tubules appeared to express PDILT (Lith, M. V., *et al*., 2007). This suggested that PDILT plays a role in sperm differentiation.

One study has suggested that PDILT can also be found in the urine. A proteomics study was undertaken to identify the differentially expressed proteins in the urine and to examine the pattern of proteomic changes occurring in the rat kidneys during progression of streptozotocin-induced model of diabetic nephropathy (DN) in rats (Sharma, V, Tikoo, K., 2014). A correlation was suggested between depleting levels of PDILT and the development of DN. Expression of PDILT was found to be highest after four weeks of induction of diabetes followed by reduction in the expression levels of PDILT during later stages. On the other hand, there are no reports showing any correlation between PDILT expression at the different stages of the disease and the development of DN. Therefore, future studies are required to confirm the expression of PDILT in specialised cells of the renal system and in the bladder, for example, and to explain the events linking the expression of PDILT with early and late stages of DN.

1.10 PDILT Structure and interactions

PDILT forms intermolecular, disulphide-dependent complexes within the testis. This was found through pulse-chase and western blot experiments (Lith, V. M., *et al.*, 2005). The post nuclear-supernatants were lysed +/- NEM to trap any disulphide bonded intermediates and were analysed by reducing and non-reducing SDS-PAGE gels. The presence of intramolecular disulphide bonds made proteins more compact and causes them to migrate more rapidly in non-reducing gels compared with reducing gels. No changes in the migration of the monomeric PDILT-Myc were observed. Overall, like PDI, PDILT engages in disulphide-bonded complexes even though the active site possesses SKQS and SKKC motifs unlike PDI (which possesses the standard redox CXXC motif).

To categorize PDILT's intermolecular disulphide interactions *in vitro*, lysates from HeLa cells were transfected with cDNA of PDILT mutants. Three PDILT mutants were made where an alanine residue replaced cysteine residues. Point mutations were made at C420, C135, and both C135 and C420 (figure 8) and were then analysed by non-reducing SDS-PAGE gel (Lith, V. M., *et al.*, 2007). When replacing either of the cysteine residues with alanine, it led to the loss of some, but not all complexes. This shows that each cysteine mediates distinct disulphide linked complexes. As expected, mutation of both cysteines abolished all disulphide-linked PDILT complexes. Overall, these cysteines are important independently in the formation of disulphide-linked PDILT complexes. However, it is unknown which proteins PDILT interacts in a cysteine dependent manner.

PDILT is known to be required for the processing and function of ADAM3 *in vivo*, a sperm: egg adhesion protein (Tokuhiro, K., *et al*., 2012). Tokuhiro, *et al.*, (2012) found that, PDILT, CALR 3 and CLGN chaperones were needed to promote ADAM

3 maturation. Co-immunoprecipitation experiments have shown that, PDILT likely interacts with CALR 3 and Calmegin directly.

Results from this group also show that CLGN directly affects ADAM 3 without the presence of PDILT. CLGN is required for ADAM 1A/ADAM 2 dimerization and subsequent maturation of ADAM 3. The effects of Calmegin disruption on the fate of Fertilin a (ADAM 1) and b (ADAM 2) during spermatogenesis were investigated by western blotting analysis (Ikawa, M., *et al.*, 2001). Experiments demonstrated the specificity of CLGN and CALR 3 chaperones. CLGN appeared to associate with ADAM 1A and ADAM 2 specifically and CALR 3 appeared to associate specifically with ADAM 3 (figure 9).

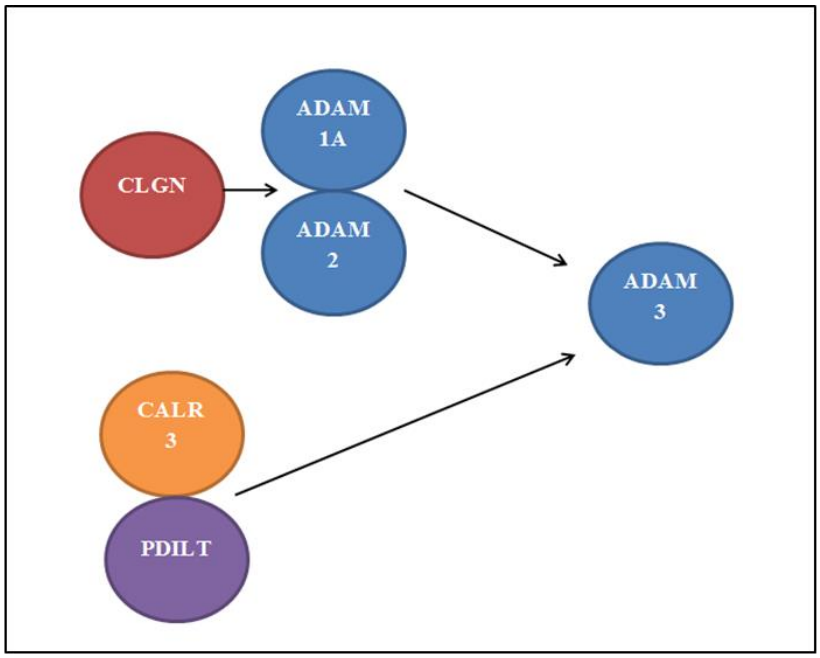


Figure 9: Interaction of chaperones involved in the folding and maturation of ADAM 3. Calmegin enables the heterodimerization of ADAM 1A and ADAM 2. CALR 3 and PDILT associate with each other, thus enabling the folding and maturation of ADAM 3. Calmegin can also associate independently with PDILT (not shown).

1.11 Calsperin (CalR 3) and Calmegin (CLGN)

As mentioned earlier in chapter 1.3, ER resident molecular chaperones aid the folding of soluble and integral membrane secreted proteins. Calnexin (CANX) and Calreticulin (CALR) proteins act as an ER lectin chaperone system which helps direct the maturation of glycoproteins in the ER (Ikawa, M., *et al.,* 2011). Calmegin (CLGN) and Calsperin (CALR 3) are exclusively expressed in the testes and are homologous to CANX and CALR. CALR 3 and CLGN are expressed in addition to CANX and CALR, so they do not replace the Calnexin cycle enzymes, but are likely to be required for specific clients in the spermatids. CLGN and CALR3 are testis-specific molecular chaperones, which are expressed in pachytene stage spermatocytes and disappear from spermatozoa upon spermination (Okabe, M., 2015). CLGN is required for the heterodimerization of ADAM 1A/ADAM 2 and this ER resident heterodimer is essential for ADAM 3 maturation (Tokuhiroa, K., *et al.,* 2012).

A set of experiments demonstrated that PDILT, CALR3 and CLGN chaperones play a role in disulphide bond formation and folding of ADAM 3 during spermatogenesis (Tokuhiro, K., *et al*., 2012). Testis lysates of mice of WT *Pdilt*, *Pdilt*$^{-/-}$, *Calr3*$^{-/-}$ and *Clgn*$^{-/-}$ mutant mice were made. Samples of these were analysed under reducing and non-reducing conditions by SDS PAGE and western blotting showed that the oxidized (folded) form of ADAM 3 was most prominent in wild type mice. However, the oxidized form was weakest in the double mutant form (*Pdilt*$^{-/-}$ and *Calr3*$^{-/-}$) as opposed to the WT and *Clgn*$^{-/-}$. On the other hand, under reducing conditions, expression levels remained the same and were unaffected by the presence or absence of the chaperones. Overall, the western blots suggest that both PDILT and

CALR 3 catalyse the formation of intramolecular disulphide bonds in ADAM 3 and therefore the folding of ADAM 3. A schematic summarising the interactions of the different chaperones together with ADAM3 is shown in figure 9.

It was found that both *Calr3*$^{-/-}$ and *Pdilt*$^{-/-}$ male mice produced apparently normal sperm but were infertile because of defective sperm migration from the uterus into the oviduct and defective binding to the zona pellucida (Ikawa, M., *et al.*, 2011, Tokuhiro, K., *et al*, 2012). Almost no *Calr3*$^{-/-}$ sperm, nor *Pdilt*$^{-/-}$ sperm were found in the uterotubal junction even 2h after copulation. Therefore, impaired sperm migration into the oviduct is the primary cause of *Calr3*$^{-/-}$ and *Pdilt*$^{-/-}$ male sterility. Sperm had defective binding to the zona pellucida *in vitro*. Thus, *in vitro* fertilization efficiency was drastically decreased in *Calr3*$^{-/-}$ sperm. Similar male infertility phenotypes were found from several other mouse knockouts including *Adam3*$^{-/-}$, *Pdilt*$^{-/-}$ and *Clmg*$^{-/-}$ animals (Ikawa, M., *et al.*, 2011; Tokuhiro, K., *et al.*, 2012).

Pdilt$^{-/-}$ male mice were infertile because ADAM 3 could not be folded properly and transported to the sperm surface without the PDILT/CALR 3 complex. However, it was later found that spermatozoa lacking ADAM3 could fertilize freshly ovulated eggs as effectively as wild-type spermatozoa when capacitated spermatozoa were directly deposited into the oviduct (Ikawa, M., *et al.*, 2001). Moreover, *Adam3*$^{-/-}$ spermatozoa successfully fertilized ZP-intact eggs *in vitro* if they were surrounded in cumulus oophorus cells. It was also reported that the *Adam1a*$^{-/-}$ spermatozoa were not able to bind to the ZP but could fertilize cumulus-rich eggs *in vitro*. In fact, it was found that primary sperm binding to the ZP could be bypassed entirely. Cumulus-conditioned media partially (but not fully) restored *Adam3*$^{-/-}$ sperm fertilizing ability to cumulus-free zona-intact eggs. Therefore, in physiological circumstances, it was suggested that the layers of cumulus oophorus may help

spermatozoa to bind/abut the ZP and induce an acrosome reaction that leads to secondary binding. Therefore, *in vivo* the PDILT/CalR 3 complex appears to play a more significant role in mobilising spermatozoa at the UTJ than in direct fertilization.

CALR 3 is a lectin-deficient chaperone directly required for ADAM 3 maturation. Endoglycosidase H treatment of ADAM 3 influenced its interaction with CLGN and CALR 3 and but did not interfere with CLGN/ADAM 3 and CALR 3/ADAM 3 interactions, demonstrating it to be lectin independent. Further examination of whether the CLGN-mediated maturation of ADAM 3 is a prerequisite for CALR 3 interaction took place, both CLGN/ADAM 3 and CALR 3/ADAM 3 interactions were not disturbed in $Calr3^{-/-}$ and $Clgn^{-/-}$ testis, respectively.

Overall, identification and understanding of key proteins in the spermatid is essential to understand the mechanism of fertilization and could potentially pave the way towards new contraception methods and treatment of male infertility.

1.12 A Disintegrin and Metalloprotease 3 protein (ADAM 3)

As mentioned previously, ADAM 1 exists in two forms in mice testis, ADAM 1A and ADAM 1B (Nishimura, et al, 2004). ADAM 1A is in the ER of testicular germ cells. However, ADAM 1B is present only in epididymal sperm on the plasma membrane. ADAM 1A and ADAM 2 (Fertilin β) form a complex to ensure the expression of ADAM 3 at the sperm cell surface. ADAM 3 is an important protein in male fertility. It is a cysteine rich, transmembrane glycoprotein, which is co-translationally translocated into the ER of spermatids during its synthesis. It is a Disintegrin and metalloprotease sperm membrane protein, which is critical for both

sperm migration from the uterus into the oviduct and for the sperm primary binding to the zona pellucida.

Initially, it had been reported that ADAM 3 was a sperm-egg fusion protein (Yuan, R., *et al*., 1997). However, it has been revealed that ADAM 3 is crucial at the stage of cell-cell adhesion with the sperm and egg but is not necessary for the fusion reaction that occurs after the stage of adhesion (Shamsadin, R., *et al.,* 1999). Using *Adam3*$^{-/-}$ knockout male mice studies, the roles of ADAM 3 have been revealed. It was found that in *Adam3*$^{-/-}$ knockout male mice, the sperm were unable to bind to the zona pellucia, again inhibiting the adhesion reaction (Yamaguchi, R., *et al.,* 2006). Lastly, it was found in *Adam3*$^{-/-}$ mutant mice that the mutant spermatozoa could not migrate through the uterotubal junction (UTJ) *in vivo* (Yamaguchi, R., *et al.,* 2006).

However, ADAM3 is a pseudogene in humans therefore other testis-specific ADAM proteins must exist and play important roles in regulating sperm adhesion and other functions. One such candidate is ADAM 5, a testis-expressed ADAM protein found in both mice and humans. Similarly, to ADAM 3, ADAM 5 is made as a precursor in the testis, processed during epididymal maturation and is located on the surface of mature sperm. Using sperm from *Adam2*$^{-/-}$ and *Adam3*$^{-/-}$ mice, it has been found that the level of ADAM 5 is modestly and severely reduced in Adam 3 and Adam 2 knockout sperm, respectively (Kim, *et al.,* 2006) Therefore, a relationship has been found between the expression of ADAM 2 and ADAM 3 with ADAM 5 implying that ADAM5 also has an important role in fertilization.

1.13 IZUMO 1 and Juno

One of the most important steps during fertilization is the fusion of the sperm and the egg. As mentioned above, there are two crucial steps to this: cell-cell adhesion and

the fusion of the sperm and the egg. The cell-cell adhesion is required for the sperm to dock onto the zona pellucida as was mentioned in the previous section in relation to ADAM 3 protein. However, in this section, the fusion of the sperm to the egg and the important proteins involved will be examined.

Juno is one of the important proteins involved in the fusion of the sperm to the egg. It is glycophosphatidylinositol (GPI) anchored protein and when unbound, it displays a globular architecture. It was originally designated as folate receptor 4 and exists on the surface of the egg. Folate receptors in general are cysteine-rich cell-surface glycoproteins that bind to folate with high affinity (Chen, C., *et al.,* 2013). However, research has found that Juno does not bind to folate (Bianchi, E., *et al*., 2014). Juno exists as a sperm-specific receptor which primarily binds to an important protein called IZUMO 1, which exists on the sperm surface. The interaction between Juno and IZUMO 1 is stable and forms a 1:1 complex (Aydin, H., *et al.,* 2016).

IZUMO 1 is involved in the fusion of the sperm to the egg and exists as a stable monomeric protein. IZUMO 1 adopts a boomerang-shaped structure and contains several domains including: a rod-shaped N-terminal four helix bundle (4HB), a hinge, and a C-terminal Ig-like domain. IZUMO 1 was shown to interact with Juno via its N-terminal domain (Aydin, H., *et al.,* 2016). All three regions of the IZUMO 1 structure interact with Juno with various different bonds including van der Waals forces, hydrophobic and aromatic interactions. When IZUMO 1 and Juno interact, a conformational change is seen within IZUMO 1 where it changes from its boomerang shape into an upright conformation. The upright conformation stabilizes the hinge region into a locked upright position, whereas when IZUMO 1 isn't bound to Juno, the hinge region exists as a dynamic flexible region.

During sperm migration, the acrosome reaction is induced before spermatozoa reach the zona pellucida and the fusion-related sperm protein IZUMO 1 on the outer acrosomal membrane is deployed to the sperm surface. Only acrosome-reacted spermatozoa can fuse with eggs. Spermatozoa without IZUMO 1 never fuse with eggs (Okabe, M., 2015). For fusion to take place, IZUMO 1 on the sperm surface must first recognize the receptor Juno which is present on the surface of the egg (Klinovska, K., *et al*., 2014). However, despite its importance recent research has suggested that IZUMO 1 doesn't exist as a direct fusion protein (Aydin, H., *et al.*, 2016). On the other hand, the conformational changes, seen when IZUMO 1 and Juno bind to each other, suggest that this adhesion triggers the progression of the IZUMO 1 4HB domain (h1, h2, h3, and h4, helices) to the vicinity of the egg membrane. These conformational changes may contribute towards the overall structural changes required for the fusion of the sperm with the egg. However, more research is needed to elaborate how IZUMO 1 and Juno molecules contribute overall to the fusion process. What is also noteworthy is that the Juno molecule is rapidly shed after the fusion process which may help to neutralize incoming acrosome-reacted sperm and prevent polyspermy from occurring (Kim, *et al.*, 2006).

1.14 Thesis Objectives

The main aim of this thesis was to further investigate the biology of PDILT. The specific objective of my thesis was to compare the expression and protein-protein interactions of PDILT mutants with wild-type PDILT. The mutants studied were PDILT-C420A (single cys mutant), PDILT- C135A, C420A (double cys mutant) and a tail mutant PDILT Δ498-580. The first aim was to observe whether mutations of certain amino acids affect the ability to form intermolecular, disulphide-depended complexes compared to wild type PDILT. The second aim for this project was to

study the interactions of PDILT with other lectin-like chaperones such as Calreticulin, Calreticulin 3 and Calmegin in mouse testis tissue. Lastly, the third aim of this thesis was to investigate the sequence relationship between PDILT and other PDIs and to assess PDILT conservation in other organisms.

2. Materials and Methods

2.1 Antibodies

Polyclonal PDI antiserum was raised in rabbits by injecting them with purified rat PDI, as described in Benham, *et al*., (2000). It was used at 1:1000 for western blotting and 1:200 for immunofluorescence. The anti-peptide PDILT anti serum was raised against purified recombinant PDILT, as described in Van Lith, *et al.*, (2007). Two bleeds from different animals were assessed, denoted αPDILT-1 and αPDILT-2. Unless otherwise stated, the αPDILT-1 serum was used. Both anti-PDILT sera were used at 1:5000 for western blotting and at 1:2000 for immunofluorescence. Monoclonal anti-myc (antigen EQKLISEEDL, Cell Signaling 9B11) was used at 1:1000 for western blotting and at 1:2000 for immunofluorescence. The calmegin, calnexin, calreticulin and calreticulin-3 antisera were kind gifts from Prof. M Okabe (The Research Institute of Microbial Diseases, Osaka, Japan) and were used at 1:1000 for western blotting. Anti-Myc was also used in immunofluorescence at 1:500 dilution and in western blotting at 1:1000 dilution. Lastly, α-RCN2 antibody was used in western blotting experiments at 1:1000. RCN2 antibody was from abcam (catalogue number ab67827)

2.2 cDNA Material

The following cDNA constructs were previously made by site directed mutagenesis using Quik Change SiteDirected Mutagenesis kit (Agilent Technologies) on pcDNA3-PDILT-myc: PDILT C420A (single Cys mutant), PDILT C135/420A (double cys mutant), PDILT 498/580 (tail mutant). The single site Cys mutant and

the double Cys mutant of PDILT have been described previously (section 1.8). The flag tagged constructs and tail deletion mutant were a kind gift from Prof. K. Nagata, Kyoto University, Japan.

Lastly, the ADAM 5 cDNA construct was expressed from a pCMV6-Myc-DDK-tagged vector. The ADAM5-myc tagged cDNA was purchased from Origene.

2.3 Cell Culture

Human fibrosarcoma HT-1080 cells were maintained in culture with Dulbecco's Modified Eagle's Medium (DMEM, Gibco, Thermofisher Scientific). The medium was supplemented with 8% (v/v) foetal bovine serum (FBS, Sigma), 2 mM glutamax (Gibco, Thermofisher Scientific), 100 units/ml penicillin and 100 µg/ml streptomycin (Sigma) at 5% CO_2. The cell line was serially passaged at 70-80% confluency, using PBS (Sigma) to wash the cells twice prior to adding 0.05% (v/v) trypsin (Gibco, Thermofisher Scientific). Passaging took place every three to four days.

2.4 Cell Lysis

Cells were lysed ~24 hours after transfection when they were ~90% confluent. The dishes were washed twice with PBS and then placed onto an ice-cold metal tray covered with ethanol-soaked tissue paper. In 6 cm dishes, 300 µl of lysis buffer (1× Triton x100, 1 µg/ml of cystatin, leupeptin, antipain, pepstatin A (CLAP) and 1× MNT (20 mM MES, 30 mM Tris-HCL, 100 mM NaCl, 1% Triton x100 (v/v)) pH 7.4 was added) 20mM NEM was placed in the lysis buffer for certain experiments to alkylate –SH groups and prevent S-S bond formation. For 10 cm dishes, 500 µl of lysis buffer was applied. The dishes were then scraped thoroughly with a cell lifter. The resulting liquid was then transferred to an Eppendorf tube and centrifuged at

16,100 g at 4˚C for 10 min and the supernatant was transferred to a fresh tube. Lysates were flash frozen in liquid nitrogen and stored at -20˚C.

2.5 Lipofectamine Transfection

Transfection was carried out using the Lipofectamine 3000 kit (Invitrogen Thermofisher). Cells that were 70-90% confluent in 6 cm dishes were washed twice with PBS.

Two mixes of the transfection reagents were made in the microbiological safety cabinet for each dish that underwent transfection:

1. 250 µl OptiMEM + 7.5 µl Lipofectamine
2. 250 µl OptiMEM + 1 µg DNA + 10 µl P3000 reagent

The dishes were then incubated for an hour at 37ºC and 5% CO_2, then 1 ml of OptiMEM was placed on the dishes and the cells were incubated overnight in a tissue culture incubator at 5% CO_2, 37˚C.

The transfection percentage efficiencies were determined by the number of cells that demonstrate perinuclear staining over the total number of cells.

2.6 Preparation of spermatid cells from mouse testis

Wild-type mouse testis samples (and PDILT knockout mice) were strain C57/BL6. All samples were taken from young adult males post-puberty (>35 days).
A mouse was dissected, then the testis was removed and decapsulated. Long segments of seminiferous tubules were removed, minced and the cellular material transferred into tubes containing PBS.

The cells were then centrifuged at 250 g and the supernatant was collected and set aside. The cell pellet was then washed and resuspended in PBS. An aliquot of cells (10 µl) were then mixed with 10 µl of trypan blue (Sigma) at 1:1 dilution in a microcentrifuge tube. The cells were observed under a haemocytometer to check the number of viable cells present.

The cells were then lysed with lysis buffer (20 mM MES, 30 mM Tris-HC, 100 mM NaCl, 1% Triton X100 (v/v), 1µg/ml CLAP, pH 7.4) on ice and centrifuged at 16,100 g at 4˚C and the supernatant was transferred to a fresh tube. Where required, the lysis buffer was supplemented with 20 mM NEM. Lysates were either used immediately for protein analysis or flash frozen in liquid nitrogen and stored at -20˚C.

2.7 Bradford Assay Protocol

The Bradford protein assay is a colorimetric assay. The protein concentrations were measured by the change in an absorbance shift of the Coomassie Brilliant Blue G-250. Under acidic conditions, the dye turned from red to blue upon binding to the protein being assayed. The binding of the protein stabilized the blue form of the Coomassie dye; therefore, the amount of the complex present in solution represented a measure for the protein concentration and was estimated by measuring the absorbance of the sample.

A 1 mg/ml stock of BSA was diluted in lysis buffer (1× Triton x100, 1µg/ml of cystatin, leupeptin, antipain, pepstatin A (CLAP) and 1× MNT (20 mM MES, 30 mM Tris-HCL, 100 mM NaCl, 1% Triton x100 (v/v), pH 7.4) to give six standard BSA concentrations of: 0, 1, 2, 4, 8, 10 µg/µl. The diluted BSA was mixed with: 0.1 M HCl, dH_2O and the Coomassie Brilliant Blue G-250 (1:4 dilution in H_2O). The

absorbance readings were measured at 595 nm in an Eppendorf Biophotometer to make the standard curve. Sample tubes were made using cell lysates, lysis buffer, 0.1 M HCl, dH_2O, and Bradford dye. The absorbance readings were then measured at 595 nm in an Eppendorf Biophotometer. Using the equation from the line of best fit of the Bradford assay standard curve, the concentration of the protein from the sample was calculated.

2.8 Immunofluorescence Protocol

Cells were grown on cover slips in 6 cm dishes overnight in complete media. In some experiments these cells had undergone transfection (Lipofectamine transfection protocol 2.5). These dishes were washed in PBS twice. The coverslips were then fixed in 4% paraformaldehyde (v/v) in PBS++ (Sigma) (PBS, 1 mM $CaCl_2$ and 0.5 mM $MgCl_2$) for 10-15 mins then washed in PBS^{++} twice. To quench the free aldehyde groups, the coverslips were then incubated in 50 mM NH_4Cl_2 for 10-15 mins and then washed in PBS^{++} 3 times. To permeabilise the cell membranes, 0.1% Triton-X100 (v/v) in PBS^{++} was added to the slips for 10 mins. The coverslips were then washed with PBS^{++} once and BSA/PBS^{++} (0.2% BSA (w/v) in PBS^{++}) twice.

The coverslips were then incubated with 25 µl of the diluted primary antibody (diluted in PBS^{++} with 0.2% BSA (w/v)) for 20 mins. Afterwards, they were washed with PBS^{++} 3 times for 5 mins each. They were then incubated with 25 µl of the fluorescent conjugated secondary antibody molecular probes, donkey anti-rabbit Alexa fluor 594 (red) at 1:1000, or goat anti-mouse Alexa fluor 488 (green) for 20-30 mins in the dark. The coverslips were then washed with BSA in PBS^{++} twice and PBS^{++} twice for 5 mins each.

The coverslips were then incubated with 40 µl of DAPI diluted to 1:1000 in PBS++ for 5 mins, enabling the cell nuclei to be stained. The coverslips were then washed in water and mounted onto microscope slides with vector shield (mounting medium). Analysis was done using a Zeiss Axiovert 200 M Apotome (OS-SIM) microscope.

2.9 SDS-PAGE

A maximum volume of 15 µl of sample was mixed 1:1 with 2x Laemmli sample loading buffer (2% SDS, 10% glycerol, 60mM Tris-HCl and 0.01% bromophenol blue). Protein prestained ladders (BioRad) were loaded onto gels to provide standards in protein molecular weight. In reducing sample buffers, the samples were also supplemented with 50 mM DTT to disrupt disulphide bonds. Samples were then heated for 5 min at 95 °C and centrifuged for 5 min at 16,100g.

The samples were analysed on 10% gels comprised of a resolving gel and a stacking gel (table 2). The resolving gel was cast in a Hoefer Dual Gel Caster for mini gels (Hoefer Inc., USA). The resolving gel was left to polymerise for 30 mins with 500 µl of water to cover the gel. The water was then removed, and the stacking gel was added along with a 10 or 15 well comb. The gel was left to polymerise for 20 mins. After polymerisation, the prepared samples were loaded next to a protein stained marker (Bio-rad). The samples were elecrophoresed at 10- 50 mA for approximately an hour in an SE250 Mighty Small II Mini Vertical Electrophoresis Unit (Hoefer) containing 1x Tris-Glycine electrophoresis buffer (10x TGS (Biorad) (25 mM Tris base, 190 mM Glycine, 0.1% SDS (v/v) pH 8.3) made up to 1x with distilled water). The gel was then either stained with Coomassie blue (Coomassie Blue staining, section 2.10) or analysed by western blotting western blotted (section 2.11).

Table 2: Composition of resolving and stacking gels

	Resolving Gel (µl)	Stacking Gel (ml)
Acrylamide 40%	4800	250
H_2O	2500	1500
1.5 M Tris pH 8.8	2500	250
SDS 10%	100	20
APS 10%	100	15
TEMED	4	2

2.10 Coomassie Blue Staining

The gel was fixed for 10 mins in 50 ml fixing solution (7% acetic acid (v/v), 40% methanol (v/v)). The solution was removed and then the Coomassie stain (80% Brilliant blue G-colloidal solution (Thermofisher) and 20% methanol (v/v)) was poured on the gel. The stain was left on the gel overnight on a rocker. The next day, the gel was destained. Firstly, it was destained in 25% methanol/10% acetic acid (v/v), for 10 mins. The solution was then poured off and a second solution of 25% methanol was then added. The gel was left in this solution on a rocker before being scanned.

2.11 Western Blotting

Following SDS-PAGE, the proteins on the gel underwent transfer. Immobilon® PVDF M membrane (Millipore) was primed in methanol for 20 seconds before placing the membrane on filter paper (towards the positive pole of the transfer apparatus) with the gel being placed on top. A sandwich was assembled using four filter papers and four sponges, moistened with transfer buffer (25 mM Tris-Base, 190 mM Glycine and 20% methanol (v/v), pH 9-9.4).

The clip holding the gel assembly was clasped shut and the assembly was placed in a tank containing an ice block and filled with transfer buffer. Transfer occurred at 150 mA for 2 hours at room temperature or at 30 V overnight at 4˚C using a Bio-rad powerpack 300.

Following transfer, the membrane was removed and blocked, protein side up, in a solution of 5% milk in TBS-T (1L: 10 mM Tris-Base, 70 mM NaCl, 1.75 mM KCl pH 9-9.4 and 0.1% Tween (v/v) pH 8) for 1 hour at room temperature or overnight at 4˚C with gentle rocking.

The membrane was then transferred to a 50 ml falcon tube and incubated with primary antibody in 1 ml of 5% milk and 2 ml of TBS-T for 1 hour with rocking before five 5-min washes in TBS-T. Secondary antibody incubation followed as for the primary, with either SARPO (anti-rabbit polyclonal) or GAMPO (anti-mouse monoclonal) at 1:3000 for 45 mins-1 hour, before five 5-min washes in TBS-T. The membrane was then dipped in TBS to remove excess detergent and transferred to Saran wrap. ECL solution (GE healthcare) was added to the membrane (500 µl, comprised of 250µl of luminol and 250µl of peroxide solution) to visualise the protein bands. The membranes were dabbed off to remove excess ECL and put in fresh saran wrap. The membranes were exposed to photographic light sensitive film (Thermofisher) in a dark room and then developed in an X-ray developer machine (XOMAT).

2.12 Immunoprecipitation

Immunoprecipitations (IP) were carried out using the required lysate and 5 µl of PDILT antibody or 5 µl calreticulin 3 antibody were immobilised on 50 µl of a 20 % suspension of Protein A or G sepharose beads (Sigma) in 150 µl of lysis buffer

(section 2.4) for 2 hours at 4 °C with rocking. Beads were spun down at 1500 g (4°C) and the lysis buffer was removed. Beads were washed in fresh 1x lysis buffer, agitated, and then washed again in the same manner. 50-100 µl of the relevant beads were added to 50 µl lysate and 150 µl lysis buffer, and rotated together for 1 hour at 4 °C. The beads were then washed five times in 500 µl of lysis buffer. Then 50 µl of the mixture (200 µl of sample buffer and 50 µl DTT)) were added to the beads. The samples were boiled at 95 °C for 3 mins on a heating block and spun down for 5 mins at 16,100 g and the supernatant was moved to a fresh tube. Thus, the eluted proteins were prepared for SDS-PAGE and Western blotting (SDS-PAGE and western blotting protocol, sections 2.9, 2.11).

3. Results

3.1 Sequence analysis and homology of PDILT

The role of PDILT in male fertility is not fully understood. In this thesis, the structure, and interactions of PDILT with other proteins were investigated to gain further understanding of the role of PDILT in the process that enables male fertility. As mentioned previously (section 1, chapter 1.8), PDILT was compared with other proteins from the PDI family to assess what structural features PDILT has which explains why it has specialised functions. PDILT was used as an input sequence in the BLAST search engine (www.ncbi.nlm.nih.gov). Comparisons were made between sequences of amino acids of PDILT protein across different species and between PDI family members in the same species. PDILT shares 27% identity (48% similarity) with PDI, however instead of the classical CXXC redox active motif, it has an SXXC motif (C420). PDILT and PDI are interesting proteins to compare as they have diverged, but share a conserved domain structure, which suggests that they are closely related.

The a and a' thioredoxin-like domains act as catalytic domains with the catalytic motif CXXC which has a major role in determining the redox potential of the enzyme (Ellgaard, L, Ruddock, L., 2005). However, there are no clear functions that have been found to be supported by the first b domain. The b' domain provides the principle protein/peptide binding site. It shares many charged and hydrophobic b' domain amino acid residues (figure 10, (Bastos-Aristizabal, *et al.*, 2014)). Since the b' domain of PDI is involved in binding peptide clients, this indicates that PDILT

has substrate-binding role or acts as a chaperone, instead of the specific function of oxidative folding. This hydrophobic pocket is conserved to some extent in PDILT.

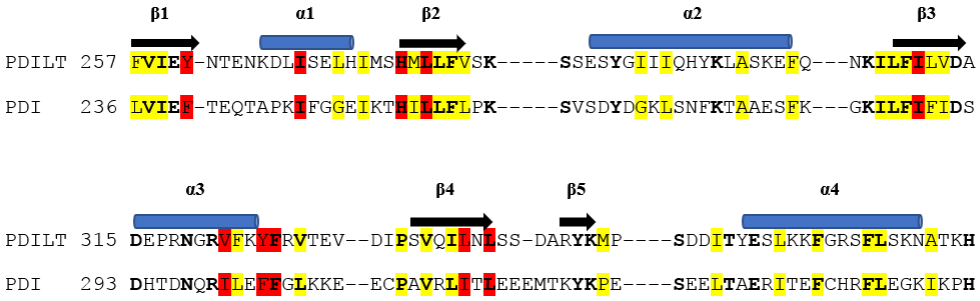


Figure 10: Structure-based sequence alignment of the b' domains of PDI and PDILT, showing alpha helices and beta strands, adapted from Bastos-Aristizabel *et al.*, (2014). Amino acids that share the same sequences are labelled in **bold**. The hydrophobic amino acids that are conserved between PDI and PDILT form are highlighted, the amino acids which form the hydrophobic substrate binding pocket in the PDILT structure and conserved between aligned proteins are highlighted in red (PDB ID 4NWY.A), which lies between the α-1 and α-3 helices.

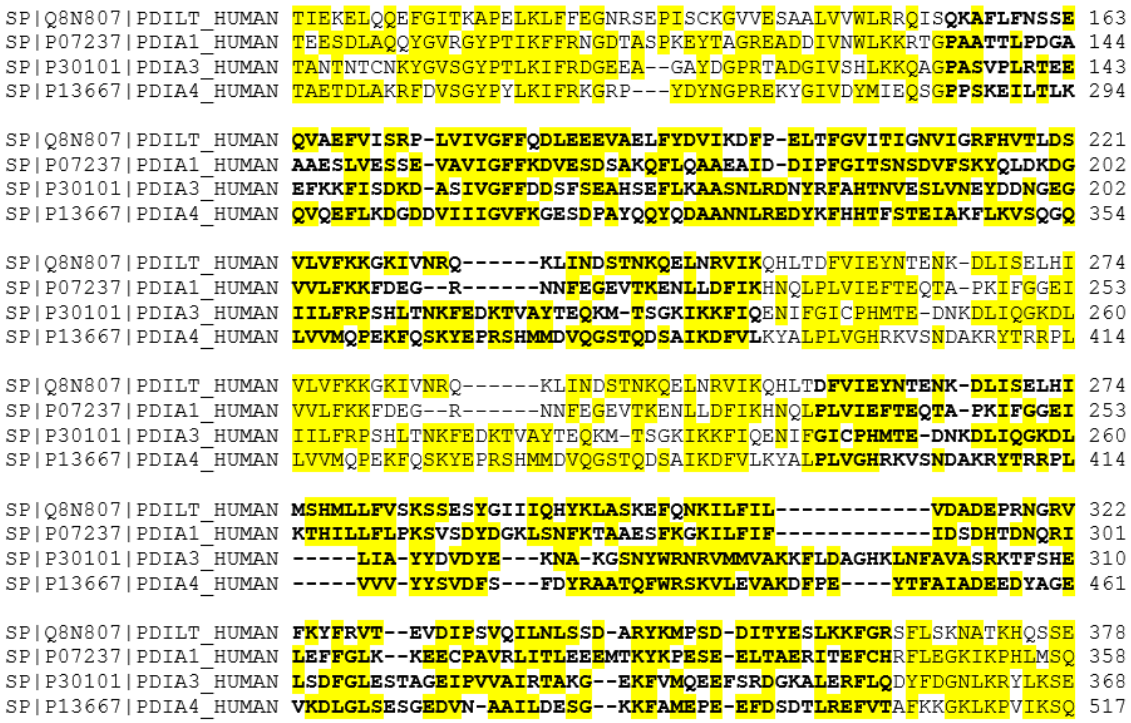


Figure 11: Combined sequence alignments of human PDI proteins. Multiple alignments of the B and B' domains of the human PDI, ERp57 and ERp72 protein. Figure 2A (in bold), shows b domain alignment of PDILT, PDIA1(PDI), PDIA3(ERp57) and PDIA4(ERp72). Figure 2B (in bold), shows b' domain alignment of PDILT, PDIA1(PDI), PDIA3(ERp57) and PDIA4(ERp72). The VLVFKK sequence is conserved in the b domain, between PDILT (amino acids: 222-227) and PDI (amino acids: 203-208). Highlighted are conserved (identical and functionally-related) residues.

PDILT also shares many conserved residues (particularly hydrophobic amino acids) with other PDI proteins including ERp57 and ERp72, as shown from an alignment between the b and b' domains of PDILT, ERp57, ERp72 and PDI (figure 11).

A comparison of the sequence of amino acids of the second catalytic domain of the PDILT protein were made across the different species. The species include *P. troglodytes* (chimp), *G. gorilla* (gorilla), *M. musculus* (mouse), *O. cuniculus* (rabbit), and *X. tropicalis* (xenopus), (figure 12a and 12b). The SXXC motif of PDILT is conserved. Chimps and gorillas share the same SKKC motif in the PDILT protein with humans. They are also identical in amino acid sequence to human PDILT.

However, mammals including *M. musculus* and *O. cuniculus* share a divergent sequence, SEKC. The *M. musculus* PDILT is 82.7% similar in amino acid identity over 588 amino acids compared with *H. sapiens* PDILT. It shares the abb'a' domain arrangement and similar characteristics such as the presence of the SXXC motif, in the a' domain. In the a domain of PDILT, *H. sapiens* has a SKQS sequence, *M. musculus* has a LKQS sequence, *O. cuniculus* has a SKQS sequence. *X. tropicalis* species has a CGHC sequence, suggesting that this protein may be redox active and not a *bone fide* redox inactive PDILT and the central two residues are not particularly important (figure 12b). In the a' domain of PDILT, *H. sapiens* has a SKKC motif, *M. musculus* s has a SEKC motif, *O. cuniculus* has a SEKC sequence and *X. tropicalis* has a SQEC motif (figure 12a). PDILT is conserved in all amphibians and fish that reproduce differently to mammals. In the a domain of PDILT, *H. sapiens, M. musculus* and *O. cuniculus* species don't contain a cysteine residue in the catalytic domain. However, the *X. tropicalis* a domain contains two cysteines residues. In the a' domain of PDILT, all the species mentioned previously, follow the general rule of having a SXXC motif, with only one cysteine present.

Using these sequences of amino acids, a family tree was generated (figure 13) which indicates the evolution of PDILT. It shows that *P. troglodytes* PDILT is closely related to *H. sapiens* PDILT as they share 97.8% similar identity of sequence of amino acids. However, *M. musculus* is only 70.1% similar in identity to *H. sapiens* PDILT and *O. cuniculus* is 79.3% similar in identity to *H. sapiens* PDILT. Therefore, *M. musculus* and *O. cuniculus* show more variation in sequence of PDILT amino acids compared with *H. sapiens* PDILT. The sequences vary more as you move away from PDILT to lower species and display lower percentage of sequence identity when compared with similar in identity compared with *H. sapiens* PDILT. For example, *X. tropicalis* PDILT is only 45.8% similar in identity compared with *H. sapiens* PDILT and *L. chalumnae* PDILT is only 33.2% similar in identity compared with *H. sapiens* PDILT.

Figure 12A: Sequence of amino acids in the region in the second catalytic domain of the a' domain of human PDILT with other species. The thioredoxin active site is labelled in red and amino acids that share the same sequence are labelled in bold. **A** demonstrates the sequence of amino acids of *H. sapiens* PDILT and *P. troglodytes* PDI protein. **B** shows the sequence of amino acids of *H. sapiens* PDILT and *G. gorilla* PDI protein. **C** demonstrates the sequence of amino acids of *H. sapiens* PDILT and *M. musculus* PDILT protein. **D** shows human PDILT and *O. cuniculus* PDI sequence of amino acids. **E** demonstrates *H. sapiens* PDILT and *X. tropicalis* PDI sequence of amino acids.

Figure 12B: Sequence of amino acids in the region in the a domain of human PDILT with other species. The thioredoxin active site is labelled in red and amino acids that share the same sequence are labelled in bold. **A** demonstrates the sequence of amino acids of *H. sapiens* PDILT and *P. troglodytes* PDI protein. **B** shows the sequence of amino acids of *H. sapiens* PDILT and *G. gorilla* PDI protein. **C** demonstrates the sequence of amino acids of *H. sapiens* PDILT and *M. musculus* PDILT protein. **D** shows human PDILT and *O. cuniculus* PDI sequence of amino acids. **E** demonstrates *H. sapiens* PDILT and *X. tropicalis* PDI sequence of amino acids.

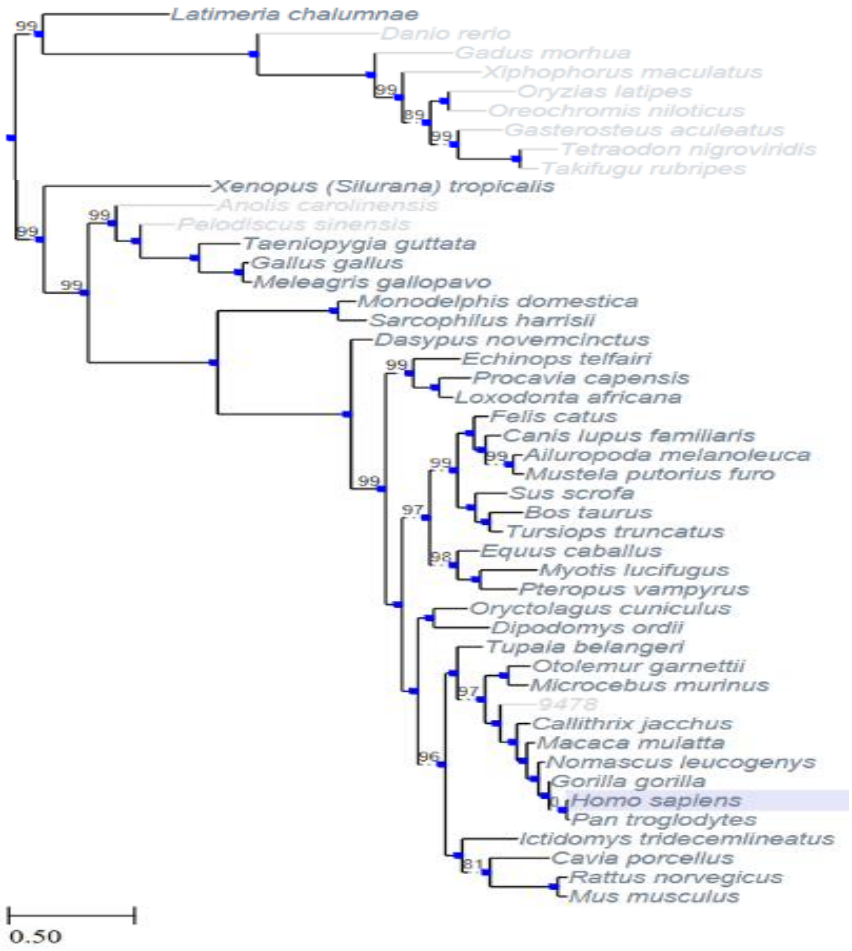


Figure 13: Family tree of different species of PDILT. The phylogenic tree was generated from comparisons of the similarity of sequences of amino acids of PDILT across different species (using EggNOG database (http://eggnogdb.embl.de). Each branch contains a "bootstrap" probability (Pb) values, which tells you how likely the relationship between the two proteins is.

Other PDI homologs also have atypical CXXC motifs. In the active site of human cells of ERp44, for example, a CRFS motif is present and unlike PDILT is conserved across other species (figure 14, appendix 1). ERp44 is a chaperone induced during ER stress and may be involved in oxidative protein folding and facilitate in the secretion of immunoglobulins (Anelli, T., *et al*, 2002).

Using the Blast search engine, the sequence of the b-domain of PDILT was compared with other PDI proteins, to see how PDILT has evolved from PDI within the PDI gene family (Figure 15). ERp27 protein was found to be closely related to PDILT. However, the function *in vivo* is largely unknown and doesn't contain a redox active CXXC motif. ERp27 contains an b and b' type domain with an ER retention sequence (LVGL) (Galligan, J. J, Petersen, D. R., 2012). PDILT is also grouped with PDI, ERp44 and ERp27. AGR2/3 proteins (necessary for mucin quality control) are grouped together and are distantly related to PDI. TMX proteins are also grouped together (ER membrane bound PDIs) and are distantly related to PDI (see appendix 4). However, results suggest there is little sequence identity. CASQ1/2 proteins are also grouped together. Calsequestrin is a calcium-binding protein of the sarcoplasmic reticulum. The protein helps to keep the calcium in the cisternae of the sarcoplasmic reticulum after a muscle contraction (Shin, D.W., *et al*., 2003)

The molecular mechanism of PDILT in facilitating ADAM3 processing to ensure male fertility is unclear. One of the ways to understand the function of PDILT is to study its interactions. Using the STRING (Search Tool for the Retrieval of Interacting Genes/Proteins, String-db.org) database, the proteins that have been

found or suggested to interact with PDILT from the literature are shown in Figure 16. These were determined from experimental procedures, as well as predicted interactions from gene fusions and gene neighbourhood experiments. The main proteins which had convincing evidence of interaction with PDILT are CLGN, CALR3, thioredoxin domain containing 11 (TXNDC11), CALR and CANX. The evidence that these proteins interact with PDILT includes co-expression, experimental/biochemical data of putative homologs, and predictions (binding and reaction). For example, PDILT-Calmegin interaction, was shown to directly co-immunoprecipitate which was found in the PNAS paper (Tokuhiro, *et al*.) The laboratory has also published some evidence of a weak interaction between PDILT and Ero1, based on co-immunoprecipititation analysis from transfected cell lines (Lith, V. M., *et al*., 2007).

The key proteins of interest, that will be the focus of this work, are the lectin-like chaperones which appear to interact with PDILT: Calreticulin 3 (CalR3), Calreticulin (CalR), Calmegin (CLGN), and Calnexin (CANX).

To study the interactions of PDILT with other proteins further, two models of expression of PDILT were established and investigated in this thesis. One approach was to use transfection of non-testis cell lines (described below in chapter 3.2.1) and the other approach was to study PDILT expression and interactions in lysates obtained from mouse testes tissue (described below in chapter 3.3).

Figure 14: Sequence of amino acids of Endoplasmic reticulum protein 44 in different species. The thioredoxin active site domain is labelled in red and amino acids that share the same sequence are labelled in bold. **A** demonstrates the sequence of amino acids of *H. sapiens* ERp44 and *P. troglodytes* ERp44 protein. **B** shows the sequence of amino acids of human ERp44 and *M. mulatta* ERp44 protein. **C** demonstrates the sequence of amino acids of human ERp44 and *C. jacchus* ERp44 protein. **D** demonstrates the sequence of amino acids of human ERp44 and *H. glaber* ERp44 protein.

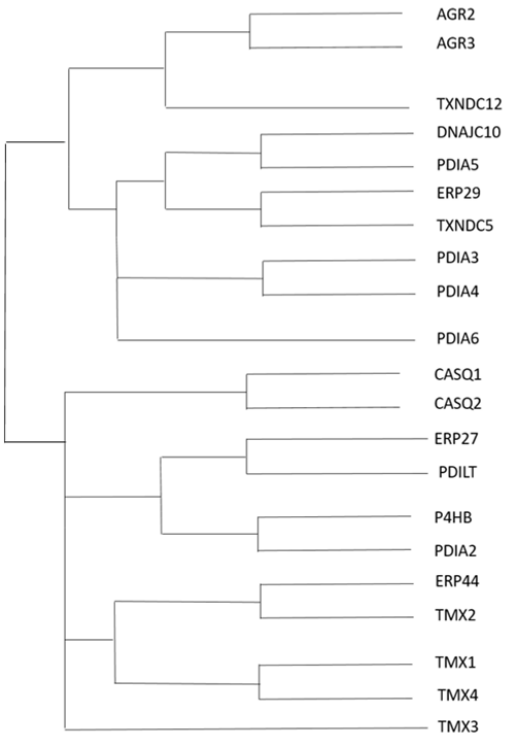


Figure 15: Phylogram of a selection of the PDI gene family proteins obtained from the NCBI BLAST search engine.

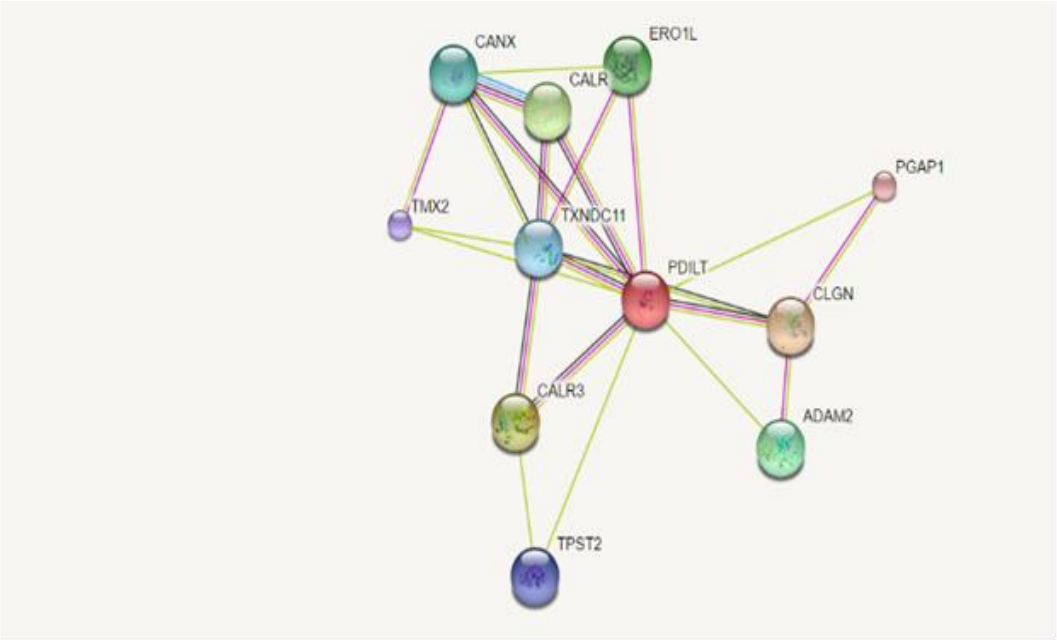


Figure 16: Protein-protein interactions between PDILT and partners. STRING database demonstrates the proteins that has been found to interact with PDILT protein. Predicted interactions are indicated if they have: green lines, by gene neighbourhood predictions, red lines, through gene fusions and dark blue lines, by gene co-occurrence. Known interactions are indicated if they have: light blue lines, which are from curated databases, and purple lines, which are experimentally determined. Other interactions also include: text mining (yellow lines), co-expression (black lines) and protein homology (blue). Known interactions are the most reliable as they will have been shown to demonstrate physical interactions through experiments such as immunoprecipitation.

3.2 Expression and interactions of PDILT in transfected HT1080 cell lines

3.2.1 Transfection of PDILT cDNA into HT1080 cells

HT1080 cells were transfected with PDILT constructs and immunofluorescence experiments were carried out to establish whether the transfections were successful and to localise the expression of PDILT (methods section, 2.8). HT1080 cells were either mock-transfected (Figure 17 A-C) or transfected with PDILT (Figure 17 D-F). HT1080 cells were stained with the primary antibodies (α-PDI and α-PDILT) and a fluorescent donkey anti-rabbit Alexa fluor 594 secondary antibody (594nm wavelength) that allowed the protein to be viewed under a red filter (figure 17B and 17E). HT1080 cells were counter-stained with DAPI to view the nucleus of the cell (figure 17A and 17D). Figure 17 shows that PDI, as expected, was resident in the ER of the cell, as demonstrated by the perinuclear staining. As also expected, PDILT was resident in the ER of the cell, as evident by a similar perinuclear staining. In Figure 17, 8 out of 16 cells were found stained in the representative field. This demonstrates that the transfection efficiency was approximately 50% for the PDILT cDNA construct.

Western blotting experiments were also undertaken to detect the presence of PDI and PDILT in the transfected HT1080 cell lysates (methods section, 2.9, 2.11). Duplicates of mock and PDILT transfected lysates were loaded onto the gel and blotted back with either the PDI antibody or the PDILT antibody. Figure 18a demonstrates that transfection was successful, as a prominent and specific band was present in the PDILT transfected lysate at around 75 kDa, as would be expected, along with the control PDI band, which is present in both the mock and the transfected lysate. Figure 18b demonstrates that cell lysis and blotting was

successful, as both lysates showed expression of PDI at around 50 kDa. Figure 18a demonstrates that transfection was successful, as a prominent and specific band was present in the PDILT transfected lysate at around 75 kDa, as would be expected, along with the control PDI band, which is present in both the mock and the transfected lysate.

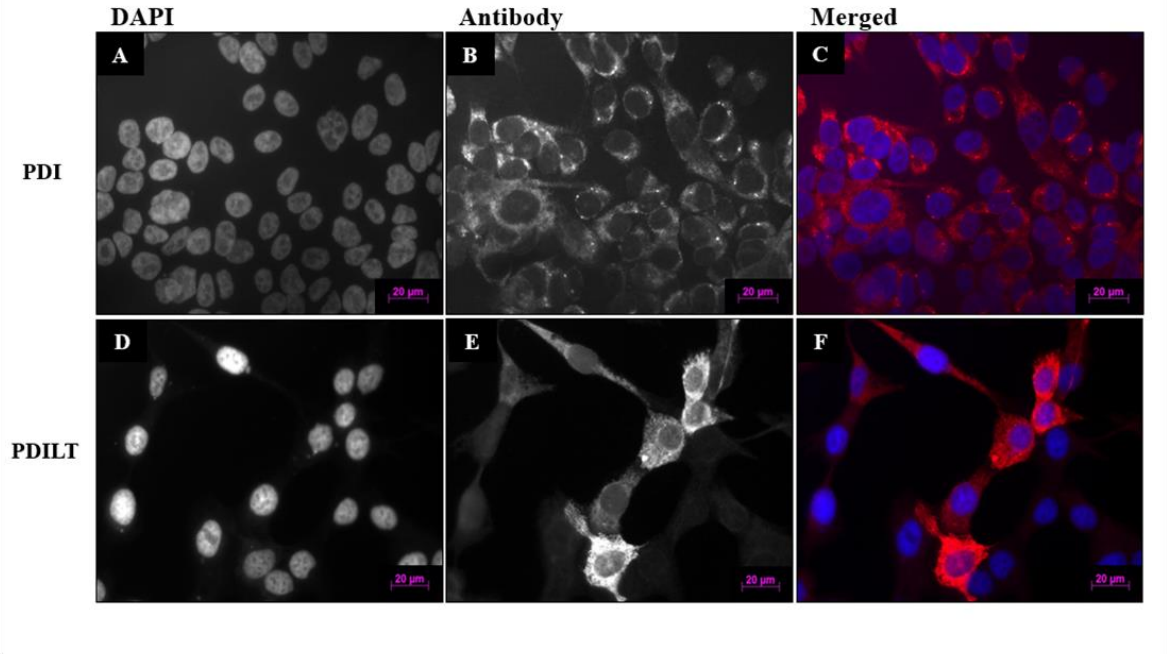


Figure 17: Immunofluorescence with α-PDI and α-PDILT. The mock transfected HT1080 cells (A-C) were stained with PDI antibody. PDI is localised in the ER as demonstrated by the perinuclear staining around the nuclei. PDILT transfected cells (D-F) were stained with PDILT antibody. PDILT is also localised in the ER as demonstrated by the perinuclear staining around the nuclei.

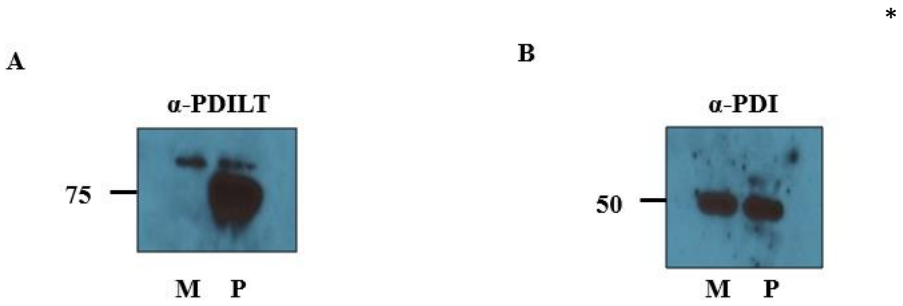


Figure 18: Western blots of Mock (M) and PDILT transfection lysates (P). Lysates of the mock transfected HT1080 cells (M) and PDILT transfected cells (P) were made and were loaded onto SDS-PAGE gels. Blot A was probed with a PDILT antibody and blot B was probed with a PDI antibody. PDILT was present at 75kDa as demonstrated in A. PDI was present at around 50kDa as demonstrated in B. *There is a non-specific band migrating at ~ 80 kDa that cross-reacts with the primary antibody and becomes visible at longer exposures.

3.2.2. Formation of disulphide-linked complexes between PDILT and partner proteins

As seen in figure 19, the non-reducing sample was first blotted for α-PDI, as a control and α-PDILT 1 and α-PDILT 2 to see which antibody performed better. Both antibodies had specificity for PDILT, the α-PDILT 1 antibody detected greater expression of PDILT from the same sample than the α-PDILT 2 sera. (Figure 19). Under non-reducing conditions, PDILT not only appeared as a band at 75 kDa but also higher molecular weight complexes were present at around 150 kDa, as expected from previous studies (Lith, V. M., 2005), demonstrating the presence of disulphide-linked complexes (Figure 19). Both antibodies could detect these complexes, but the reactivity was much stronger with the anti-PDILT1 sera.

In figure 20, reducing samples (boiled in the presence of DTT in the sample buffer) and non-reducing samples (without the presence of DTT) were loaded and blotted back with α-PDILT 1. Under non-reducing conditions (without the presence of DTT), there was a band at 75 kDa, as seen previously, which demonstrates the presence of PDILT (Figure 20). There were also a couple of higher molecular weight bands indicating the presence of disulphide-linked complexes like those seen in Figure 20. However, on the same gel under reducing conditions (presence of DTT) there was only one band present at 75kDa and there were no higher molecular weight complexes (Figure 20). There appears to be less protein in the reduced sample compared with the non-reducing sample when exposed for 10 minutes. Although, when using a lighter exposure (for 5 minutes) you can still see the presence of the higher molecular weight bands in the non-reducing sample (see appendix 2).

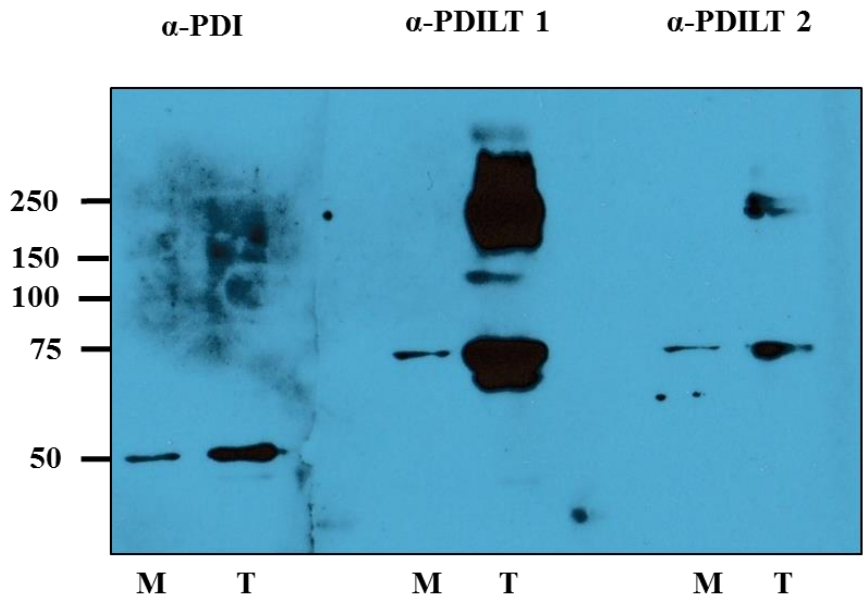


Figure 19: Western blot of Mock and PDILT lysates under non-reducing (NR) conditions. The two lysates were blotted for PDI antibody, both lysates displayed expression of PDI. Mock lysate (HT1080 cells only) acted as a control compared with the PDILT transfection lysate. When blotting for PDILT, the transfection lysate (more so when using PDILT 1 antibody) displayed two thick bands, indicating expression of PDILT. There were also higher molecular weight bands present indicating the presence of disulphide-bonded complexes.

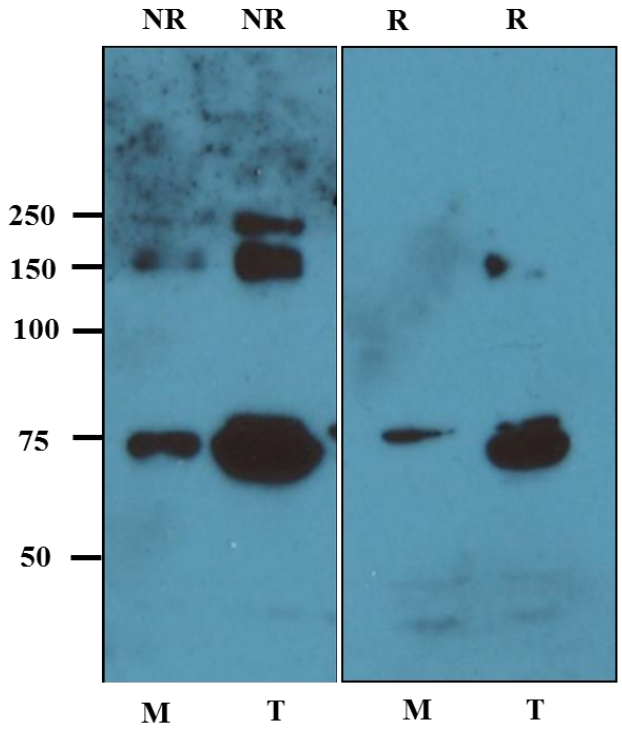


Figure 20: Western blot of Mock and PDILT lysates under reducing and non-reducing conditions. The two lysates were blotted with α-PDILT1 with and without the presence of DTT (reducing and non-reducing conditions). Mock lysate (HT1080 cells only) act as a control to compare with the PDILT transfection lysate which displays a thicker band, demonstrating the expression of PDILT. Under non-reducing conditions, higher molecular weight complexes were found. However, under reducing conditions these complexes disappeared.

Figure 21: Domain organisation of PDILT and position of the mutations. The a and a' domains are shown in red, the b and b' domains are shown in blue, and the tail domain is shown in orange. The KEEL sequence enables ER retention and the signal sequence targets the protein to the ER.

3.2.3. Expression of PDILT WT and mutants of PDILT: PDILT C420A (single Cys mutant), PDILT C135A/C420A (double Cys mutant) and PDILT Δ498-580 (tail mutant)

The importance of cysteine residues 135 and 420 in the formation of inter-molecular disulphide bonds and the importance of the tail domain for the integrity of PDILT protein were investigated. To investigate the role of the cysteine residues and the tail domain, HT1080 cells were transfected with wt PDILT cDNA, PDILT C420A (single Cys mutant), PDILT C135AC420A (double Cys mutant) and PDILT Δ498-580 (tail mutant) for immunofluorescence and western blotting experiments (Figure 21). To establish the localisation of the PDILT mutants compared with wt PDILT, immunofluorescence experiments were undertaken. As seen in figure 22, transfection appears to besuccessful, as perinuclear staining around the nucleus was evident in all cells transfected with the mutant PDILT constructs. However, future work should include co-staining with another ER protein to act as a control, so we can be confident that the localisation was not affected by these mutations. All mutants demonstrate ER localisation therefore none of the residues that were mutated are essential for ER localisation. However, transfection efficiency was somewhat lower than previous transfections from previous experiments. The cause of this maybe because expression of certain PDILT proteins are toxic and therefore fewer positive/weaker stained cells were detected. The transfection efficiency for the following constructs were: wild type PDILT cDNA approximately 40%, single cys mutant PDILT cDNA 50%, double cys mutant PDILT cDNA 35%, and the tail mutant PDILT cDNA approximately 60% efficiency.

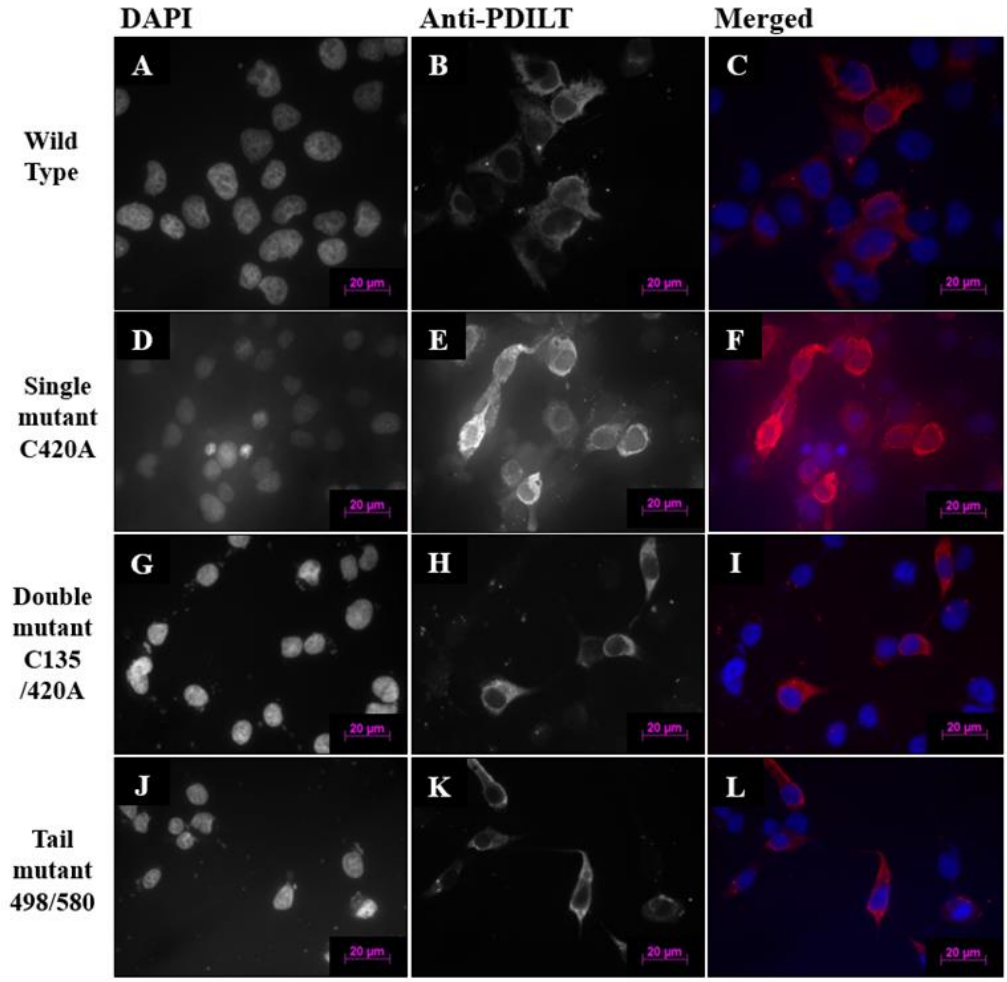


Figure 22: Immunofluorescence of HT1080 cells transfected with WT PDILT cDNA and PDILT mutants' cDNA which is stained with α-PDILT. Figure A-C demonstrates the successful transfection of Wild Type PDILT cDNA. Previous experiments have shown that PDILT is localised in the ER as demonstrated by the perinuclear staining around the nuclei. Figures D-F demonstrate the transfection of single mutant C420A PDILT cDNA. Figures G-I demonstrates the transfection of double mutant C135A/420A. Figures J-L shows the transfection of tail mutant 498/580. Figures D-I demonstrate the typical perinuclear staining around the nuclei but it is not as clear as seen in the wild type.

To visualise the proteins in the lysates, the samples were loaded onto a 10% SDS PAGE gel, electrophoresed and the gel was stained with Coomassie blue (figure 23). The gel demonstrated an even loading of proteins from the mock (HT1080 cells without transfection), WT (PDILT), single cys mutant (PDILT C420A), and tail mutant (PDILT Δ498-580). However, there appears to be less protein loaded in the double cys mutant (PDILT C135AC420A).

To further understand the behaviour of the PDILT C420A, C135AC420A mutants and the tail mutant, transfected HT1080 cell lysates were analysed by western blot. Samples made from the lysates were analysed under both reducing and non-reducing conditions (i.e. with/without the presence of 50 mM DTT). The samples were blotted back for α-PDI as a control. As is evident in figure 24, all lysates expressed PDI protein, as bands were present at around 50 kDa in all the lysates.

To see the effects of the site-specific mutations on PDILT, protein lysates were blotted back for α-PDILT (figure 25). Under reducing conditions, bands appear at ~75 kDa, as expected for WT PDILT. As expected, no bands appeared from Mock lysate (HT1080 cells without transfection), as PDILT was not present. Both single cys mutant and the double cys mutant lysates showed that PDILT migrated at 75 kDa, demonstrating that these mutations did not influence the stability/degradation of PDILT under steady state conditions. However, in the tail mutant, a faint band appeared lower down the gel at approximately 65 kDa, consistent with the reduction in molecular weight caused by truncating the tail. The lower expression level suggested that loss of the tail mutant might influence the stability of the PDILT protein. The c-terminal tail of PDILT also contains a positively charged region (KKKTSEEVVVVVAKPKGPPVQKKKPK), which contains a stretch of hydrophobic amino acids (as depicted in red), which may be required (possibly with

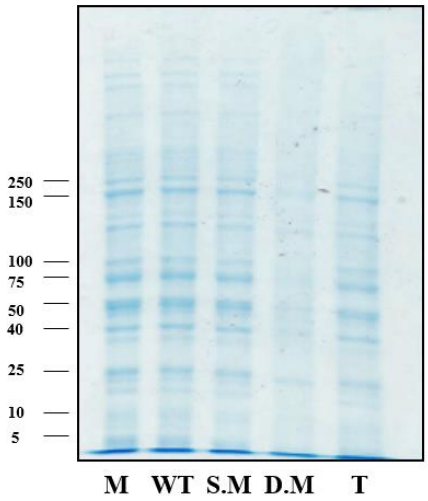


Figure 23: Coomassie stained SDS-PA gel of PDILT WT and mutant lysates. The lysates are: (M) Mock (HT1080 cells), (WT) Wild Type (PDILT Transfection), (S.M.) single mutant (transfection of PDILT with mutation on C420A), (D.M.) double mutant transfection of PDILT with mutations on C135A/420A), (T) tail mutant (transfection of PDILT with mutations on 498/580). They each demonstrate a relatively even amount of expression of protein.

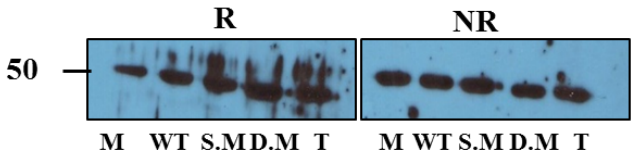


Figure 24: Western blot of PDILT WT and mutant lysates, under reducing and non-reducing conditions. Blotted for α-PDI. (M) Mock (HT1080 cells), (WT) Wild Type (PDILT Transfection), (S.M.) single mutant (transfection of PDILT C420A), (D.M.) double mutant (transfection of PDILT C420A), (T) tail mutant (PDILT delta498-580). PDI was equally recovered in each transfectant, indicating successful lysis and blotting.

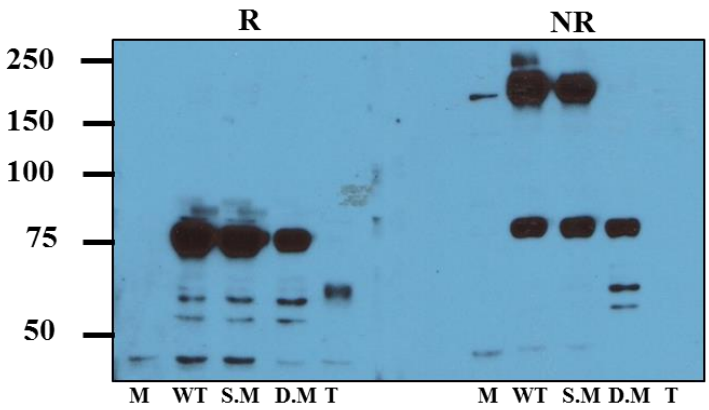


Figure 25: Western blot of the PDILT WT and mutant lysates, under reducing and non-reducing conditions. Blotted for α-PDILT. (M) Mock (HT1080 cells), (WT) Wild Type (PDILT Transfection), (S.M.) single mutant (transfection of PDILT C420A), (D.M.) double mutant (transfection of PDILT C420A), (T) tail mutant (PDILT delta498-580). Under reducing conditions WT, S.M., and D.M. PDILT was expressed at 75kDa. The tail mutant was expressed at a lower molecular weight but the band was faint. Under non-reducing conditions, higher molecular weight complexes were present in the W.T. and the S.M. lysate but disappeared in the D.M. lysate. The tail mutant lysate had no band present on the gel (the film was exposed for 2 mins). The mock lysates showed no expression of PDILT, acting as a negative control.

other domains) for interactions with proteins such as CLGN and CALR3. To study this further a co-expression with CLGN and CALR3 would be needed followed by immunoprecipitation experiments. Under non-reducing conditions, bands again appeared at 75kDa for the wild type (PDILT), the single cys mutant (PDILT C420A) and the double cys mutant (PDILT C135A/C420A) lysates, but not for the tail mutant (PDILT Δ498-580). As expected no bands were detected in the mock lysate (HT1080 cells without transfection), as there was no PDILT present. Higher molecular weight complexes were also present in wild type, PDILT lane and in PDILT C420A mutant transfection lane. However, the higher molecular weight complexes disappeared in PDILT C135A/C420A and in PDILT Δ498-580. Therefore, the tail domain and the two cysteine residues C135A/C420A are important for the formation of disulphide-linked complexes in PDILT. However, one cysteine residue (C135A) alone is enough to form disulphide-linked complexes.

The experiment shown in Figure 25 was repeated and the film was exposed for a longer period to see if the tail mutant band was present (figure 26). After a longer exposure, the tail mutant appeared at lower molecular weight to that of the wild type (as expected) and in the double mutant, PDILT protein again did not form higher molecular weight complexes (see appendix 3 for other exposures of figure 26).

To delve further into the effects of PDILT mutants on protein-protein interactions, immunoprecipitation experiments were carried out on the PDILT WT and mutants to study the interactions of PDILT with the RCN2 protein. Reticulocalbin 2 (RCN2) protein is a calcium-binding protein located in the lumen of the ER; previous experiments in the laboratory have suggested that PDILT might interact with this protein (unpublished observations). The beads from the immunoprecipitation experiments were incubated with the PDILT antibody and the immunoprecipitates

were run on SDS-PAGE and then probed on a western blot with an RCN2 antibody. The size of RCN2 is 43kDa. Unfortunately, the band size is close to that of cross-reactive antibody, so it was difficult to establish whether RCN2 co-immunoprecipitated with PDILT, although the tail mutant, in lane IP T lane, IP experiment appeared to reveal a band at 43kDa (figure 27).

To establish a positive control for RCN2 expression, fibroblast cells, HT1080 cells and mouse testis tissue was used to make lysates. These three lysates were then loaded onto a western blot and were blotted back for α-PDI in Figure 28A as a control for lysis and protein transfer. Figure 28B shows a western blot of the three lysates immunoblotted with α-RCN2. Only fibroblast cells appear to show expression of the RCN 2 protein, suggesting that RCN2 is not expressed in HT1080 cells. To explore potential interactions between RCN2 and PDILT further, co-immunoprecipitation of PDILT and RCN2 from testis could be carried out, or fibroblasts could be transfected with PDILT to see if endogenous RCN2 co-immunoprecipitates with the transfected protein.

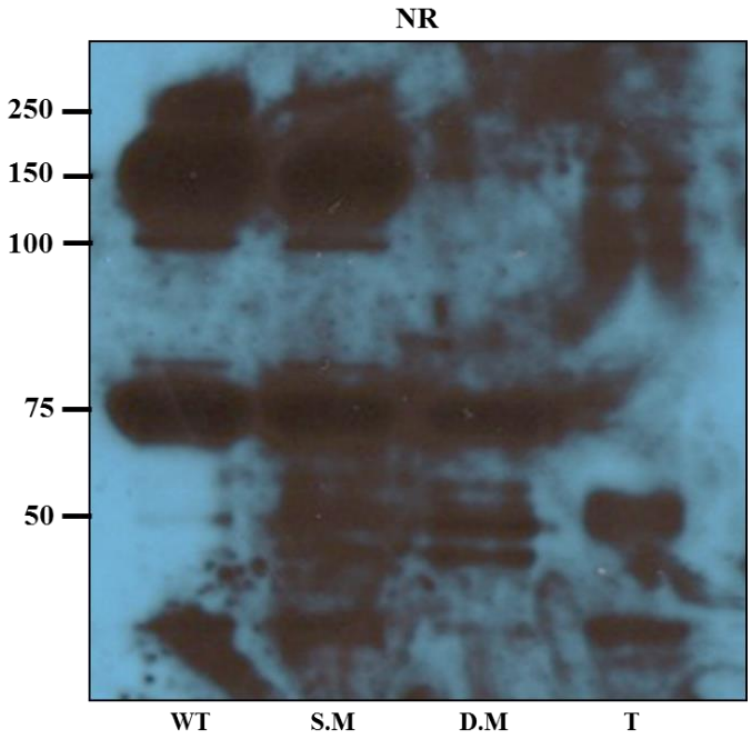


Figure 26: Western blot of the PDILT WT and mutant lysates, under non-reducing conditions. Blotted for α-PDILT. (WT) Wild Type (PDILT Transfection), (S.M.) Single Mutant (transfection of PDILT with mutation on C420A), (D.M.) Double Mutant transfection of PDILT with mutations on C135A/420A), (T) Tail Mutant (transfection of PDILT with mutations on 498/580). Under non-reducing conditions, higher molecular weight complexes are present in the W.T. and the S.M. but like figure 12 the higher molecular weight complexes disappear in the D.M. lysate. The tail mutant lysate expresses a stronger band at around 55kDa (the film was exposed for 10 minutes). This indicates that the PDILT was expressed in the Tail Mutant lysate but at a lower expression.

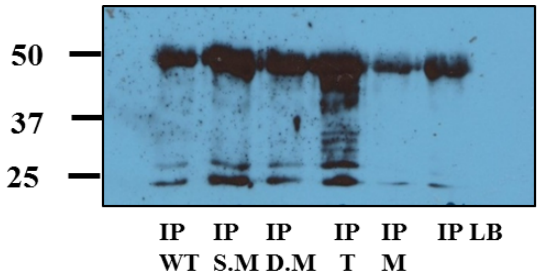


Figure 27: Western blots of the Immunoprecipitation experiments of the PDILT WT and mutant lysates, blot back with RCN2 antibody. IP mock acts a control, only one band is present showing the presence of the beads and antibody. IP lysis buffer (LB) also acts as a control which again only shows one band demonstrating the presence of the beads and the antibody. The band that shows the interaction of PDILT with RCN2 is close in molecular weight to the antibody.

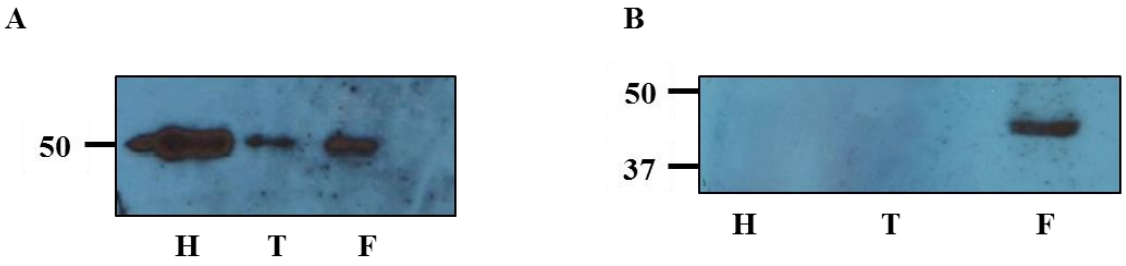


Figure 28: Western blots of HT1080 cells lysate, testis tissue lysate and fibroblast cell lysate. Figure A has the three lysates loaded onto a gel and blotted for α-PDI as a control. Figure 18B shows the three lysates once again loaded onto a gel and blotted with α-RCN2 but only fibroblast cells expressed RCN2.

3.2.4 ADAM 5/Myc and PDILT

As described in the introduction, ADAM 3 is an important protein in male fertility. Previous findings have suggested that the mature ADAM 3 is involved in the initial binding of the sperm to the zona pellucida on the egg. However, ADAM 3 is not processed properly in the absence of PDILT, so a model system is required. ADAM 3 is a pseudogene in humans, therefore other testis-specific in humans, additional ADAM proteins are likely to play important roles in regulating sperm functions. One candidate is ADAM5, which is made as a precursor in the testis, is processed during epidydimal maturation and is located on the sperm surface in mice. It would be interesting to see if ADAM 5 interacts with PDILT to facilitate investigations into the role of PDILT in human ADAM protein processing and quality control.

To investigate whether ADAM 5 interacts with PDILT a transfection model was set up. HT1080 cells were either mock-transfected, transfected with PDILT, or transfected with ADAM 5 cDNA. To determine whether transfection was successful, immunofluorescence experiments were done, as shown in Figure 29. Figure 29 A-C shows a control of mock-transfected HT1080 cells immunostained with the PDI antibody. Figure 29 D-F demonstrates transfection of PDILT, although the transfection efficiency is low (approximately 40%). Lastly, Figure 29 G-I shows a transfection of ADAM 5; however, expression of ADAM 5 could not be conclusively demonstrated because the ADAM5-expressing cells appear to be unhealthy.

To check protein expression, cell lysates were loaded onto a western blot and blotted back for α-PDI, α-PDILT and α-Myc (figure 30). As a control, PDI expression was confirmed at ~50 kDa. PDILT was successfully expressed, as evidenced by a prominent band present at around 75 kDa in the PDILT transfected cells. However,

there was no band present when blotted for α-Myc where it would be expected, at around 57 kDa. This is likely due to the death of transfected cells. Therefore, the transfection of ADAM 5 wasn't successful. Further optimisation of ADAM5 transfection and expression in a range of transfectable cell lines will be required to establish whether ADAM5 is a potential client for PDILT.

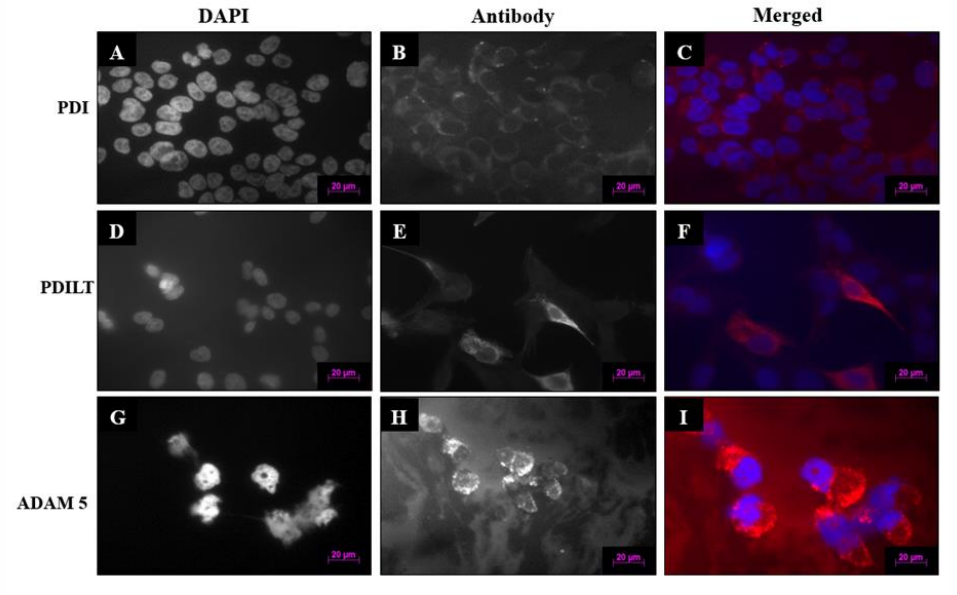


Figure 29: Immunofluorescence of Mock, PDILT and ADAM 5 protein. Figure A-C demonstrates a control of HT1080 cells stained with α-PDI. The control was established as the PDI was, as seen previously, localised in the ER as demonstrated by the perinuclear staining around the nuclei. Figures D-F demonstrates the transfection of PDILT cDNA, which was stained with α-PDILT as some cells had perinuclear staining as seen in the control. Figures G-I demonstrates the transfection of ADAM 5 protein, which was stained with α-Myc. Again, some cells demonstrated transfection, as evident by the perinuclear staining but the cells were unhealthy compared to other experiments.

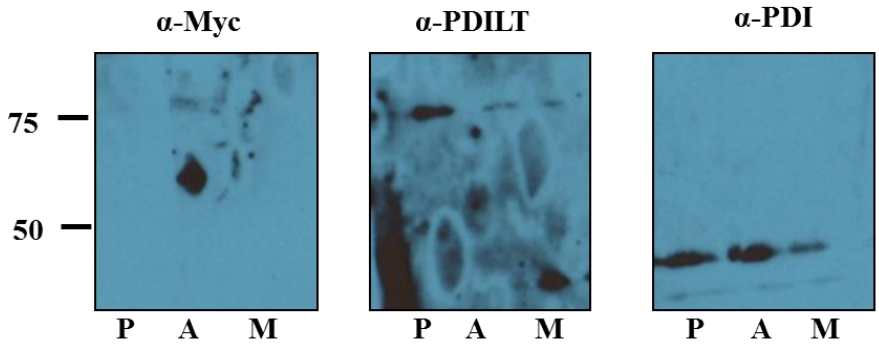


Figure 30: Western blot of ADAM 5 transfected lysate (A), PDILT transfected lysate (P) and mock lysate – HT1080 cells only (M). They lysates were blotted with α-Myc, α-PDILT, and α-PDI. There was expression of PDI protein, indicating lysis was successful and PDILT was present only in P cells. However, ADAM 5 was not expressed so transfection of ADAM 5 was unsuccessful.

3.3. PDILT interactions in mouse testis tissue

3.3.1. Expression of PDILT in mouse testis lysates

To study the interactions of PDILT with chaperone proteins, it is preferable to study them *in vivo*. Therefore, mouse testes were dissected, and lysates were made from the preparation of spermatid cells (methods, 2.7), with and without the presence of NEM to trap inter-molecular disulphide bonds. Cells were counted using a haemocytometer, with an average preparation giving ~6 x10^8 cells. Using a Bradford assay, the protein concentration of the lysates was measured. The protein concentration of the testis lysate with the presence of NEM was 8.13 mg/ml and the protein concentration of the testis lysate without the presence of NEM was 7.94 mg/ml in the example shown.

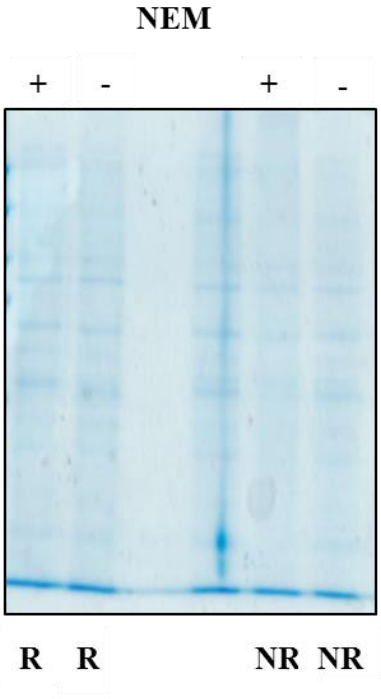


Figure 31: Coomassie staining of the mouse testis tissue lysates (lysed with/without the presence of NEM). Samples of the lysates (with and without the presence of NEM) were made under reducing (R) and non-reducing (NR) conditions.

Samples of the lysates were then made under reducing and non-reducing conditions. These were loaded onto an SDS-PAGE gel and stained with Coomassie blue (figure 31) to confirm equal loading and sample recovery.

3.3.2. Detection of PDILT and lectin-like chaperone proteins

To establish whether the PDILT protein engages in higher molecular weight complexes under non-reducing conditions, as found previously from transfection analysis, mouse testis lysate (under reducing and non-reducing conditions), were loaded onto a gel and blotted back for PDILT (figure 32), alongside transfected HT1080 cell lysate samples (T). In the presence of NEM and absence of DTT (reducing agent), higher molecular weight complexes appear to form in both lysates, indicating the formation of disulphide bond complexes. On the other hand, under reducing conditions (presence of DTT) and absence of NEM no higher molecular complexes formed apart from a small amount in the reducing lysate lane, perhaps indicating that the sample wasn't fully reduced. Overall, these findings are consistent with previous findings that disulphide linked complexes are present in PDILT, *in vivo* and *in vitro* (figure 32, Lith, V. M., 2005). However, the quality of the western blot compromises the reliability of the data. Therefore, ideally this experiment would need to be replicated in future experiments.

Another observation was that NEM appeared to trap intermolecular disulphide bonds complexes in both samples, as under non-reducing conditions there were no higher molecular weight complexes present. However, with the presence of NEM higher molecular weight complexes formed.

Next, three chaperone proteins that are likely to interact with PDILT and play an important role in the function of PDILT and its role in male fertility were assessed,

namely CALR, CLGN, CALR3 (figure 33). Samples from mouse testis tissue and PDILT-transfected lysates were analysed by SDS-PAGE and blotted back for PDILT, again under non-reducing conditions; higher molecular weight disulphide bond complexes were present at around 150 kDa. Under reducing conditions, the samples were also blotted back with CALR antibody. A band was detected at around 55 kDa, indicating expression of CALR in both samples. These samples were also blotted back with the CANX antibody, two strong bands were present in both samples at approximately 90kDa, indicating the expression of CANX.

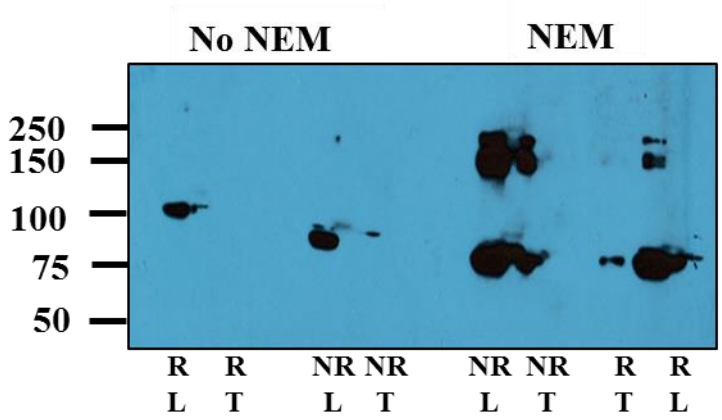


Figure 32: Western Blot of mouse testis (L) lysates (lysed with/without the presence of NEM) and lysate HT1080 cells transfected with PDILT cDNA (T). Samples of these lysates were made under reducing (R) and non-reducing (NR) conditions. These were blotted with α-PDILT. When blotting for PDILT, the transfection bands were faint but the mouse testis tissue lysate bands were well defined. Under non-reducing conditions, the molecular weight bands appeared more strongly in the L lysates than in the T lysates.

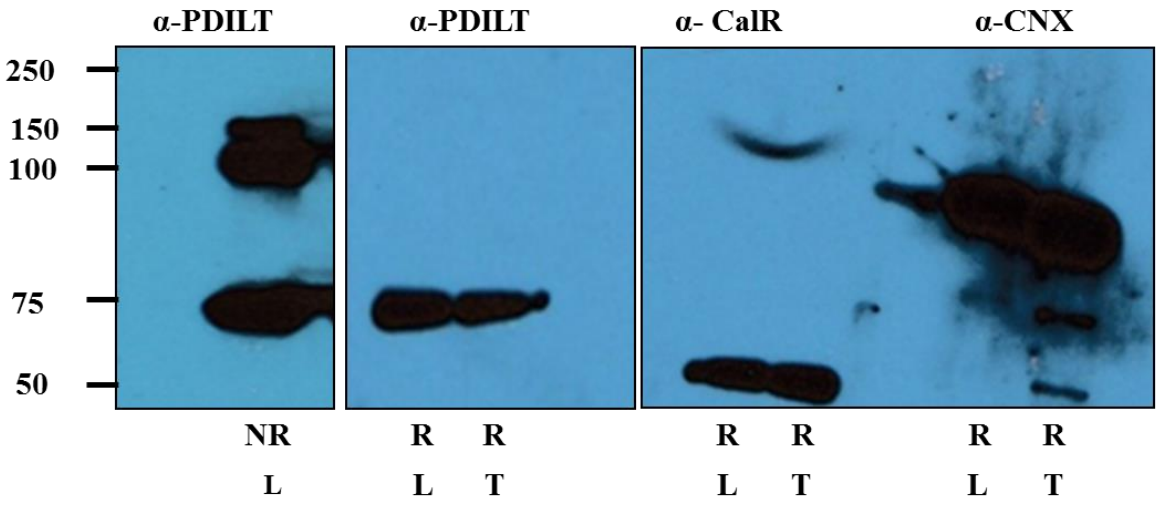


Figure 33: Western Blot of mouse testis (L) lysate and lysate HT1080 cells transfected with PDILT cDNA (T). Samples were made under reducing (R) and non-reducing (NR) conditions. The samples were blotted with α-PDILT, α-Calreticulin and α-Calmegin antibodies. Expression of these proteins was found in testis lysates.

3.3.3. Interactions of PDILT with the lectin-like chaperone proteins CALR, CALR3, and CLGN

To study whether PDILT interacts with chaperones in the mouse testis, immunoprecipitation experiments were undertaken (methods section, 2.12).

In Figure 34 A, as a positive control, the protein A sepharose beads were incubated with PDILT antibody which was washed and incubated with the mouse testis lysate, then blotted back with the PDILT antibody in a western blot. In the IP + L lane, the antibody heavy chain was detected at ~50kDa, as seen in the control (IP + LB) but there was also another band at 75 kDa as seen in the lysate, indicating that PDILT was immunoprecipitated by the beads. In the IP + LB there is only a 50kDa band present, which establishes a positive antibody control.

In Figure 34 B, the beads were incubated with the PDILT antibody and incubated with the mouse testis lysate. After SDS-PAGE the immunoprecipitate was blotted back with the CALR antibody in a western blot. The antibody heavy chain's molecular weight (50kDa) is close to the molecular weight of the CALR protein (55kDa). Therefore, it was difficult to establish whether the immunoprecipitation was successful. However, comparing the lanes of the IP + LB and the lysate, they seemed to show separation in molecular weight which establishes good controls. Also, the anti-CALR band (IP+L) appeared thicker, which may indicate that PDILT is interacting with CALR in the lysate, as they are likely to contain both the CALR antibody heavy chain band and the band indicating the presence of CALR. Crosslinking experiments will be required to elucidate whether CALR can interact weakly with PDILT.

Protein A sepharose beads were incubated with PDILT antibody, which was washed and then incubated with the mouse testis lysate, and analysed for the presence of

CLGN using the CLGN antibody in a western blot (Figure 34 C). The IP + LB lane indicates a positive control of the CLGN antibody heavy chain and in the L lanes there is a band showing at around 93kDa which indicates the molecular weight the protein should be at. The appearance of multiple antibody bands in the IP + L lanes made it difficult to see if the immunoprecipitation was successful as the expected molecular weight of the CLGN protein is between 75-100 kDa. However, there appeared to be a faint band in the IP + L sample isolated in the absence of NEM, indicating there may be immunoprecipitation of CLGN, as suggested by previous publications (e.g. Lith, V. M., *et al*., 2007). Therefore, this evidence suggests that PDILT may interact with CLGN in the lysate. Lastly, the presence of NEM also appears to diminish the interaction of PDILT with CLGN.

Finally, Protein A sepharose beads were incubated with the CALR3 antibody, washed, and then incubated with the mouse testis lysate prior to analysis by SDS-PAGE and western blotting (Figure 34 D). The immunoprecipitates were blotted back with the PDILT antibody in a western blot. In all four of the IP + L samples (with and without the presence of NEM treatment), there was a distinct band at approximately 75kDa, which was also seen in the lysate. There was also a band at approximately 50kDa where the antibody lies. Overall, the data indicates that CALR3 interacts with PDILT strongly in the mouse testis.

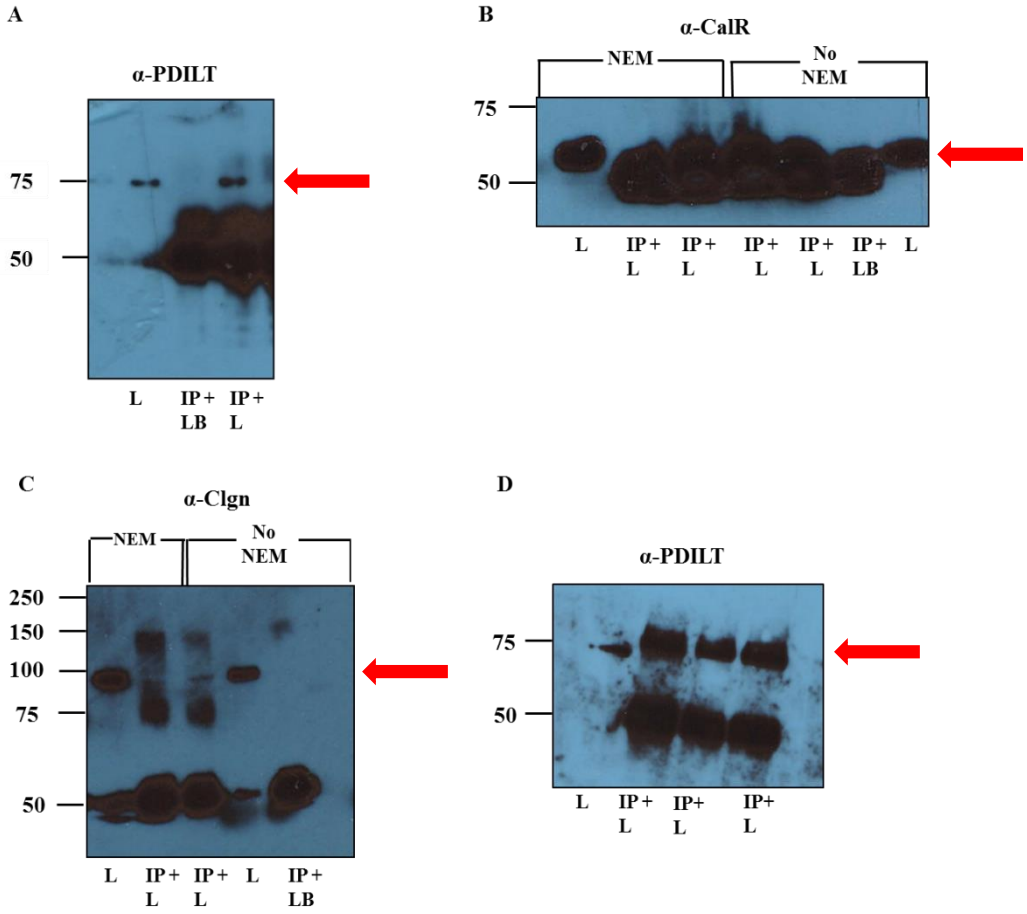


Figure 34: Western Blot of IP experiments of the mouse testis lysates (lysed with/without the presence of NEM) and to act as controls there is: lysate on its own and IP lysis buffer (LB). The red arrows indicate the molecular weight the protein of interest should be. The IP + LB indicates only one band at 50kDa which show the antibody's heavy chain. A) The lysate was incubated with α-PDILT and blotted back for α-PDILT as a control. PDILT was detected at 75kDa (as where the arrow indicates) demonstrating the presence of PDILT. B) The lysate was incubated with α-PDILT and blotted back for α-Calreticulin. Although the bands are not distinct from each other, there appears to be the presence of Calreticulin protein (at 63kDa by the red arrow). This suggests that PDILT and Calreticulin were interacting. C) the lysate was incubated with α-PDILT and blotted back for α-Calmegin. Although faint, there appears to be band at around 93kDa which would indicate that Calmegin was present and interacting with PDILT. D) the lysate was incubated with α-Calreticulin 3 and blotted back for α-PDILT. There are distinct bands at 75kDa indicating that Calreticulin 3 was interacting with PDILT.

4. Discussion

4.1 Structural conservation of PDILT

In this thesis, sequence analysis was made of the four domains of PDILT: a, b, b' a'. In most PDIs, the a and a' thioredoxin-like domains act as catalytic domains which catalyse redox reactions (Lith, M. V., *et al*. 2007). However, the a' domain consistently has an SXXC motif across many species in PDI proteins, including: *P. troglodytes, G. gorilla, M. musculus, O. cuniculus* and *X. tropicalis* (chapter 3.1, figure 12). This suggests that the SXXC motif is important in the function of the a' thioredoxin-like domain in PDILT proteins (Lith, V. M., 2005). The b' domain of PDILT has a protein/peptide binding site and is likely involved in substrate-binding role, as it contains a hydrophobic pocket (Lith, V. M., 2005). This hydrophobic pocket is also consistently seen in PDI across species, as well as in the ERp57 and ERp72 proteins. There are no clear functions from the PDILT b domain, but it also contains hydrophobic amino acids, similar to the PDI, ERp57 and ERp72 proteins. Therefore, the b domain may also be involved in the binding to other proteins.

4.2 The roles of PDILT and its functions

The roles of PDILT in male fertility are not fully understood. However, evidence has suggested that PDILT engages in disulphide dependent interactions in the ER and interacts with other proteins (Lith, V. M., 2005). The molecular mechanism of PDILT in male fertility hasn't been well established. Therefore, in this thesis, PDILT was studied by establishing a transfection model in HT1080 cells and by analysing mouse testis tissue lysates to study PDILT protein-protein interactions *in vitro* and *in*

vivo. Lastly, through sequence analysis the sequence conservation of PDILT was assessed.

In transfected cell lines, the specific localisation of PDILT in the ER was confirmed using immunofluorescence experiments (chapter 3.2.1, figure 17). The ER localisation of PDILT and a range of mutants supports the previous findings (Lith, V. M., 2005), as seen in chapter 3.2.3 and demonstrates that PDILT expression is tolerated by non-testicular cell types. The expression pattern of PDILT seems to extend much further from the perinuclear region than that of PDI, which is more punctate. The successful transfection of HT1080 cells with PDILT cDNA was also further analysed by western blotting of the lysates where PDILT was detected in the transfected lysates (figure 18).

In both the transfected cell line model and in the mouse testis tissue lysate model, the formation of disulphide-linked complexes in PDILT has been established (section 3.2.2. and figure 19-20). Under non-reducing conditions, higher molecular weight complexes were found when western blotting was used to analyse PDILT, in both the transfected cell lysates and the mouse testis tissue lysates. This indicates the presence of intermolecular disulphide bonds within PDILT. Furthermore, in reducing gels the higher molecular weight complexes disappear. This supports the idea that the higher molecular weight bands represent the intermolecular disulphide bonds because under reducing conditions samples are treated with dithiothreitol (DTT). DTT is a reducing agent, which most likely reduces the disulphide bonds in PDILT to form dithiols (figure 20). This also supports previous findings (Lith, V. M., *et al.* 2005), as described in the introduction, chapter 1.10.

Another finding which suggests that intermolecular disulphide bonded complexes are formed between PDILT and other proteins comes from the mouse testis tissue lysates. During lysis, the lysates were treated also with and without N-ethylmaleimide (NEM), then samples were then further treated with/without DTT. The samples were then blotted back for PDILT (figure 25). NEM appeared to trap intermolecular disulphide bonds. Under non-reducing conditions and with the presence of NEM, there were higher molecular weight complexes present. However, under reducing conditions (with the presence of DTT) without the presence of NEM, no higher molecular weight complexes formed. NEM is an alkene which is reactive towards thiol groups and modifies cysteine residues. Therefore, the presence of NEM would trap the disulphide bonds, which supports the findings found in figure 32, where the higher molecular weight complexes were formed in the presence of NEM.

To further delve into the structural properties of PDILT, point mutations were made of the cDNA of PDILT. Along with the WT cDNA, these mutated cDNAs were transfected into HT1080 cells and lysates were made (section 3.2.3). These amino acids were studied because they are of keen structural interest. The two cysteine amino acids were mutated to study the importance of these amino acids in the formation of disulphide bonds. The tail was also studied, as little is known about this structure, to see the effects of the structure of PDILT when the domain is compromised.

It was found that both WT PDILT and the C135A and C420A mutants, both formed higher molecular weight complexes. However, there appears to be a loss of some higher molecular weight complexes but still a strong band indicating that the monomer was still present. As expected, the C135A/C420A double mutation caused

the disappearance of the higher molecular weight complexes indicating their role in inter-molecular disulphide bond formation. This confirms the findings found in a previous paper (Lith. M. V., *et al*., 2007). Interestingly, in the Δ498-580 tail mutant, PDILT was detected less efficiently than the C135A, C420A and wt PDILT proteins. Higher exposures of film were required to detect the presence of bands. However, a band was located at a lower molecular weight than the expected 75kDa. Faint bands were also present at higher molecular weight indicating there may be higher molecular weight complexes present but not as extensive as seen in the WT nor in the single cysteine mutant. Therefore, the tail may be an important domain in stabilising the structural integrity of PDILT. For future experiments, these experiments and transfections should be replicated. Pulse chase experiments will also be required to follow up this observation and establish the half-life of PDILT in the absence of the tail domain.

4.3 Interactions of PDILT with chaperone proteins: CALR, CALR3, and CLGN

To study the interactions of PDILT with other proteins, the mouse testis tissue was analysed using western blotting. CALR and CNX (chapter 3.3.2, figure 33) were detected by co-immunoprecipitation when blotting with anti-CALR and CLGN antibodies. To see whether these chaperone proteins directly interact with PDILT chaperone protein, immunoprecipitation experiments were made on the mouse testis tissue lysate.

As mentioned in section 3.3.3, immunoprecipitation experiments were performed to study the interactions with PDILT and other proteins such as: CALR, CLGN and CALR3. From the immunoprecipitation experiments PDILT appears to be

interacting with CALR protein in the mouse testis tissue lysate. However, the molecular weight of the antibody heavy chain and the CALR antibody are very similar therefore it is difficult to be sure but it is consistent with the findings found from the STRING database (String-db.org) suggesting that these proteins interact. Immunoprecipitation experiments also suggest that CLGN interacts with PDILT, however it wasn't clear due to the presence of multiple antibody bands near the CLGN protein band. On the other hand, the interaction of PDILT with CLGN supports the findings found from the STRING (String-db.org) database that these proteins interact. Also, the presence of NEM appeared to diminish the interaction of PDILT with CLGN, which is interesting as NEM traps disulphide bonds. This suggests that the formation of disulphide bonds is important in the interaction of PDILT with CLGN.

Lastly, CALR3 has been successfully shown to interact strongly with PDILT, which again supports previous findings from the group.

4.4 Future Directions

A lot more is now known about PDILT, particularly the a' domain and its SXXC motif. It is known to form intermolecular disulphide-dependent complexes and likely interacts with other chaperones, such as CALR3. However, it is not known which amino acids in what parts of the PDILT protein interact with other chaperones such as CALR3. This knowledge would help to build a clear picture of the role of PDILT in the mechanism in male fertility. One interesting finding was that NEM may interfere with the interaction of PDILT and CLGN. Therefore, it would be good to analyse further immunoprecipitation experiments with CLGN mutants and PDILT with and without NEM. This would elucidate if NEM indeed interferes with a direct

interaction of these proteins and PDILT cys mutant analysis would confirm whether any cys residue is required for the interaction of these proteins.

The tail domain was found to be of structural importance to PDILT, so it may be involved in the folding of the protein and maintaining its structure. The exact role of the tail domain remains to be elucidated but it may be worth investigating the half-life of PDILT in the absence of PDILT. It may also be worth performing specific sequence analysis of the tail, to see if there are hydrophobic amino acids present, to understand the structure of PDILT.

Another important area of research is the specific interactions of PDILT with chaperones. The results obtained suggested that PDILT interacts with CALR, CALR3 and CLGN. However, the potential interaction seen between PDILT and CALR needs to be explored further, as previous findings (chapter 3.1, figure 16 STRING database, String-db.org) only show findings of these interactions with putative homologs in other species, i.e. not in *H. sapiens.* Therefore, more research is needed to confirm this interaction in *H. sapiens* lysates. The function of CALR is unknown, however, we can speculate that CALR may interact with a specialist chaperone, such as CALR3 and aid it in the interaction of PDILT, as CALR3 was shown to interact with PDILT directly. Another potential function of CALR is that PDILT may require it for its own quality control. Lastly, little is known about the interactions between CLGN and PDILT, therefore more research is needed to elucidate the interactions between the two.

Steps were taken to create an ADAM5 client model, which is made as a precursor in the testis. ADAM 5 cDNA was transfected in HT1080 cells and the transfection efficiency was detected using immunofluorescence experiments and western

blotting. This system would be useful to study in the future for the interactions of ADAM 5 protein with PDILT by immunoprecipitation and western blotting. It would have also been interesting to study the interactions of these chaperones with the PDILT mutants to understand which amino acids are crucial for the interactions with ADAM 5. This would provide a clearer picture of the role of PDILT in the mechanism involved in enabling male fertility.

4.5 Conclusions

This thesis has explored many avenues in elucidating the structure and function of PDILT protein.

It has explored the evolution of the structure of PDILT. Analysis of hydrophobic pocket amino acids, and how they are conserved across species and between other PDI homologs, demonstrates the importance of certain amino acids (e.g. SXXC motif in the a' domain, and the hydrophobic amino acids in the b and b' domains). This allows us to gain a clearer picture of its important features and how these specialised domains allow PDILT to carry out its functions (e.g. the a and a' domains catalyse redox reactions, and the b and b' domains are likely involved in the protein/peptide binding or substrate binding roles).

It has elucidated the importance of disulphide bonds in the interactions with other chaperone proteins. For instance, the formation of disulphide bonds was important in the interaction of PDILT with CLGN. It also pinpointed that C135A and C420A amino acids in the PDILT protein are important for inter-molecular disulphide bond formation. Through point mutations of the cDNA of PDILT, a proposed function of the tail domain was made in that it may play an important role in the stability of the structural integrity of PDILT.

It has also uncovered interactions of different chaperones, such as CALR3 with PDILT. There also appeared to be interactions of PDILT with CALR, this may be to assist with the direct interactions of PDILT with CALR3, or to enable quality control of PDILT. Lastly, PDILT appeared to interact with CLGN with the aid of disulphide bonds. With further experiments, such as immunoprecipitation of more chaperones, more immunoprecipitation experiments with the same chaperones but using treatment of NEM at different concentrations to monitor the effects of NEM in the interactions of these chaperones with PDILT and lastly, pulse chase experiments of PDILT with and without the presence of the tail domain to establish the half-life of PDILT. This would clarify importance of disulphide bonds in the interaction of PDILT with other chaperones and would confirm the importance of the tail domain for the integrity of PDILT protein.

The mapping of PDILT interacting proteins would be another step towards elucidating the mechanism of action of this unusual chaperone, and will help further understand the mechanism of sperm-egg binding and sperm-UTJ adhesion during reproduction of PDILT. Lastly, this thesis began to develop a working ADAM 5 client model. More research is needed to establish an ADAM 5 client model for the future study of the interactions with PDILT. This would provide the bigger picture of the role of the PDILT protein in the mechanism of male fertility. The elucidation of these mechanisms could help towards developing new therapies for male infertility and could also aid in researching new forms of contraception for men.

5. References

Agarwal, A, Mulgund, A, Hamada, A, Chyatte, M. R., 2015. A unique view on male infertility around the globe. Reproductive Biology and Endocrinology, 13(37). DOI: 10.1186/s12958-015-0032-1

Alberti, A, Karamessinis, P, Peroulis, M, Kyperou, K, Kavvadas, P, Pagakis, S, Politis, P. K, Charonis, A., 2009. ERp46 is reduced by high glucose and regulates insulin content in pancreatic beta-cells. Am J Physiol Endocrinol Metab, 297(3):E812-21. DOI: 10.1152/ajpendo.00053.2009

Amman, R. P., 2008. The cycle of seminiferous epithelium in humans: a need to revisit? Journal of Andrology, 29(5): 469-87. DOI: 10.2164/jandrol.107.004655

Anelli, T, Alessio, M, Mezghrani, A, Simmen. T, Talamo, F, Bachi, A, Sitia, R., 2002. ERp44, a novel endoplasmic reticulum folding assistant of the thioredoxin family. EMBO J, 21(4): 835-844. DOI: 10.1093/emboj/21.4.835

Appenzeller-Herzog, C, Ellgaard, L., 2008. The human PDI family: Versatility packed into a single fold. Biochimica et Biophysica Acta 1783(4) 535-548. DOI: 10.1016/j.bbamcr.2007.11.010

Aydin, H, Sultana, A, Li, S., Thavalingam, A., Lee, J. E., 2016. Molecular architecture of the human sperm and egg JUNO fertilization complex. Nature Letter 534, 562-565. DOI:10.1038/nature18595

Bastos-Aristizabal, S, Kozlov, G, Gehring, K., 2014. Structure of the substrate-binding b' domain of the Protein disulphide isomerase-like protein of the testis. Scientific Reports 4 (4464). DOI: 10.1038/srep04464

Behnke, J, Hendershot, L. M., 2013. The large Hsp90 Grp170 binds to unfolded protein substrates *in vivo* with a regulation distinct from conventional

Hsp70s. The Journal of Biological Chemistry 289, 2899-2907. DOI: 10.1074/jbc.M113.507491

Benham, A. M., 2012. The protein disulfide isomerase family: key players in health and disease. Antioxidants & Redox Signaling 16(8):781-9. DOI: 10.1089/ars.2011.4439

Berruti. G, Paiardi. C., 2011. Acrosome biogenesis. Spermatogenesis, 1(2): 95-98. DOI: 10.4161/spmg.1.2.16820

Bianchi, E, Doe, B, Goulding, D, Wright, G. J., 2014. Juno is the egg Izumo receptor and is essential for mammalian fertilization. Nature 508, (7497): 483-7. DOI: 10.1038

Bulleid, N. J., 2012. Disulfide Bond Formation in the Mammalian Endoplasmic Reticulum. Cold Spring Harb Perspect Biol 4(11). DOI: 10.1101/cshperspect.a013219

Chen, C, Ke, J, Zhou, E, Yi, W, Brunzelle, J. S, Li, J, Yong, E. L, Xu, H. E, Melcher, K., 2013. Structural basis for molecular recognition of folic acid by folate receptors. Nature Letter 500, 486-489. DOI: 10.1038/nature12327

Cooper, G. M., 2000. The Cell: A Molecular Approach. 2nd edition. Sunderland (MA): Sinauer Associates. The Endoplasmic Reticulum. [ONLINE] https://www.ncbi.nlm.nih.gov/books/NBK9889/

EggNOG, The EggNOG datatbase team, 2016, http://eggnogdb.embl.de/#/app/results#ENOG411CXKD_datamenu

Egozcue, S, Blanco, J, Vendrell, J. M, Garcia, F, Veiga, A, Aran, B, Barri, P. N, Vidal, F, and Egozcue, J., 2000. Human male infertility: chromosome anomalies meiotic disorders, abnormal spermatozoa and recurrent abortion. Hum. Reprod. 6, 93– 105

Ellgaard, L, Helenius, A., 2003. Quality control in the endoplasmic reticulum. Nat Rev Mol Cell Biol 4(3): 181-91. DOI: 10.1038/nrm1052

Ellgaard, L, Molinari, M, Helenius, A., 1999. Setting the standards: quality control in the secretory pathway. Science 286 (5446) 1882-1888. DOI: 10.1126/science.286.5446.1882

Ellgaard, L, Ruddock, L. W., 2005. The human protein disulphide isomerase family: substrate interactions and functional properties. EMBO Rep. 6(1):28-32. DOI: 10.1038/sj.embor.7400311

Frickel, E. M, Frei, P, Bouvier, M, Stafford, W. F, Helenius, A, Glockshuber, R, Ellgaard, L., 2004. Erp57 is a multifunctional thiol-disulfide oxidoreductase. J Biol Chem. 279(18):18277-87. DOI: 10.1074/jbc.M314089200

Galligan, J. J, Petersen, D. R., 2012. The human protein disulfide isomerase gene family. Hum Genomics. 6(1): 6. DOI: 10.1186/1479-7364-6-6

Green, N. P. O, Stout, G. W, Taylor, D. J., 1985. Biological Science 2. Pg 749-751. ISBN: 0 521 26951 2

Hatahet, F, Ruddock, L.W., 2009. Protein Disulfide Isomerase: A Critical Evaluation of Its Function in Disulfide Bond Formation. Antioxidants & Redox Signaling 11(11): 2807-50. DOI: 10.1089/ars.2009.2466

Hargreave, T. B., 2000. Genetic basis of male fertility. Br. Med. Bull. 56(3): 650-71. PMID: 11255552

Ikawa, M, Nakanishi, T, Yamada, S, Wada, I, Kominami, K, Tanaka, H, Nozaki, M, Nishimune, Y, Okabe, M., 2001. Calmegin Is Required for Fertilin α/β Heterodimerization and Sperm Fertility. Developmental Biology 240(1): 254-261. DOI: 10.1006/dbio.2001.0462

Ikawa, M, Tokuhiro, K, Yamaguchi, R, Benham, A. M, Tamura, T, Wada, I, Satouh, Y, Inoue, N, Okabe, M., 2011. Calsperin Is a Testis-specific Chaperone Required for Sperm Fertility. J. Biol. Chem. 286(7):5639-46. DOI: 10.1074/jbc.M110.140152

Inaba, K., 2011. Sperm flagella: comparative and phylogenetic perspectives of protein components. Molecular Human Reproduction 17(8):524-38. DOI: 10.1093/molehr/gar034

Inoue, N, Ikawa, M, Isotani, A, Okabe, M., 2005. The immunoglobulin superfamily protein Izumo is required for sperm to fuse with eggs. Nature 434(7030):234-8. DOI: 10.1038/nature03362

Jankovicova, J, Simon, M, Antalikova, J, Cupperova, P, Michalkova, K., 2014. Role of tetraspanin CD9 molecule in fertilization of mammals. Physiol. Res. 64: 279-293. ISSN: 1802-9973

Kim, E, Baba, D, Kimura, M, Yamashita, M, Kashiwabara, S, Baba, T., 2005. Identification of a hyaluronidase, Hyal5, involved in penetration of mouse sperm through cumulus mass. PNAS 102(50):18028-33. DOI: 10.1073/pnas.0506825102

Kim, T, Oh, J, Woo, J. M, Choi, E, Im, S. H, Yoo, Y. J, Kim, D. H, Nishimura, H, Cho, C., 2006. Expression and relationship of male reproductive ADAMs in Mouse. Biology of Reproduction 74(4):744-50. DOI: 10.1095/biolreprod.105.048892

Klinovska, K, Sebkova. N, Dvorakova-Hortova. K., 2014. Sperm-Egg Fusion: A molecular enigma mammalian reproduction. Int. J. Mol. Sci. 15: 10652-10668. DOI: 10.3390/ijms150610652.

Kohno, K, Normington, K, Sambrook, J, Gething, M. J, Mori, K., 1993. The

promoter region of the Yeast KAR2 (BiP) Gene contains a regulatory domain that responds to the Presence of Unfolded Proteins in the Endoplasmic Reticulum. Molecular and Cellular Biology, 13(2) 877-890. PMCID: PMC358971

Lith, V. M, Karala, A.-R, Bown, D, Gatehouse, J. A, Ruddock, L.W, Saunders, P. T. K, Benham, A. M., 2007. A Developmentally Regulated Chaperone Complex for the Endoplasmic Reticulum of Male Haploid Germ Cells. Mol. Biol. Cell 18(8): 2795-2804. DOI: 10.1091/mbc.E07-02-0147

Lith, V. M, Hartigan, N, Hatch, J, Benham, A. M., 2005. PDILT, a Divergent Testis-specific Protein Disulfide Isomerase with a Non-classical SXXC Motif That Engages in Disulfide-dependent Interactions in the Endoplasmic Reticulum.The Journal of Biological Chemistry, 280(2): 1376-83. DOI: 10.1074/jbc.M408651200.

Li, Z, Srivastava, P., 2004. Heat-shock proteins. Curr Protoc Immunol. Appendix 1: Appendix 1T. DOI: 10.1002/0471142735.ima01ts58

Martin-Du Pan, R. C, Bischof, P, Campana, A, Morabia, A., 1997. Relationship between etiological factors and total motile sperm count in 350 infertile patients, Archives of Andrology, 39:3, 197-210, DOI: 10.3109/01485019708987917.

Marzec, M, Eletto, D, Argon, Y., 2012. GRP94: an HSP90-like protein specialized for protein folding and quality control in the Endoplasmic Reticulum. Biochim Biophys Acta, 1823(3): 774-787. DOI: 10.1016/j.bbamcr.2011.10.013. PMCID: PMC3443595

Miki, K., 2007. Energy metabolism and sperm function. 1 Soc Reprod Fertil Suppl. 65:309-25. PMID: 17644971

NCBI, Gene, CALR3 calreticulin 3 [*Homo sapiens* (human)] [ONLINE] Available at http://www.ncbi.nlm.nih.gov/gene/125972

Nishimura, H, Kim, E, Nakanishi, T, Baba, T., 2004. Possible function of the ADAM1a/ADAM2 Fertilin Complex in the appearance of ADAM3 on the sperm surface. J Biol Chem. 13;279(33):34957-62. DOI: 10.1074/jbc.M314249200

Nishikawa. S, Brodsky. J. L, Nakatsukasa. K., 2005. Roles of molecular chaperones in Endoplasmic Reticulum quality control and ER-associated degradation (ERAD). J. Biochem. 137(5):551-5. DOI: 10.1093/jb/mvi068

Nixon-Abell, J, Obara, C. J, Weigel, A. V, Li, D, Legant, W. R, Xu, C. S, Pasolli, H, A, Harvey, K, Hess, H. F, Betzig, E, Blackstone, C, Lippincott-Schwartz, J., 2016. Increased spatiotemporal resolution reveals highly dynamic dense tubular matrices in the peripheral ER. Science. 354, 6311. DOI: 10.1126/science.aaf3928

Okabe, M., 2015. Mechanisms of fertilization elucidated by gene-manipulated animals. Asian Journal of Andrology 17(4)646-652. DOI: 10.4103/1008-682X.153299

Onda, Y, Onda, Y., 2013. Oxidative Protein-Folding Systems in Plant Cells, Oxidative Protein-Folding Systems in Plant Cells. International Journal of Cell Biology, International Journal of Cell Biology. 2013(2013) DOI: 10.1155/2013/585431.

Ortega, C, Verheyen, G, Raick, D, Camus, M, Devroey, P, Tournaye, H., (2011). "Absolute asthenozoospermia and ICSI: What are the options?". *Human Reproduction Update*. **17** (5): 684–692. doi:10.1093/humupd/dmr018. PMID 21816768.

Peckham, M, Knibbs, A, Paxton, S., 2003. The Histology guide, Male reproduction. http://www.histology.leeds.ac.uk/male/

Rajfer, J., 2003. Decreased levels of testosterone in the aging male. Urology. 5(Suppl 1): S1–S2. PMCID: PMC1502317.

Riemer, J, Hansen, H. G, Appenzeller-Herzog, C, Johansson, L, Ellgaard, L., 2011. Identification of the PDI-Family Member ERp90 as an Interaction Partner of ERFAD. PLoS ONE, 6(2): e17037. DOI: 10.1371/journal.pone.0017037

Schlegel, P. N., 2004. Causes of azoospermia and their management. Reprod. Fertil. Dev, 16(5): 561-72. PMID: 15367371 DOI: 10.10371/RD03087

Sharma, V, Tikoo, K, 2014. Stage-specific quantitative changes in renal and urinary proteome during the progression and development of streptozotocin-induced diabetic nephropathy in rats. Mol Cell Biochem 388(1-2):95-111. DOI: 10.1007/s11010-013-1902-5

Shamsadin, R, Adham, I. M, Nayernia, K, Heinlein, U. A. O, Oberwinkler, H, Engel, W., 1999. Male mice deficient for germ-cell Cyritestin are infertile. Biology of Reproduction. 61(6):1445-51. PMID: 10569988

Shin, D. W, Pan, Z, Kim, E. K, Lee, J. M. Bhat, M. B, Parness, J, Kim, D. H, Ma, J., 2003. A Retrograde Signal from Calsequestrin for the Regulation of Store-operated Ca^{2+} Entry in Skeletal Muscle. JBC 278, 3286-3292. DOI 10.1074/jbc.M209045200

Simoni, M, Weinbauer G. F, Gromoll, J, Nieschlag, E., 1999. Role of FSH in male gonadal function. Ann Endocrinol (Paris) 60(2):102-6. PMID: 10456180

Sitia, R, Braakman, I., 2003. Quality control in the endoplasmic reticulum protein factory. Nature 426(6968):891-4. DOI: 10.1038/nature02262

Stockton, J. D, Merkert, M. C, Kellaris, K. V., 2003. A Complex of Chaperones and

Disulfide Isomerases Occludes the Cytosolic Face of the Translocation Protein Sec61p and Affects Translocation of the Prion Protein. Biochemistry, 42(44): 12821-34. DOI: 10.1021/bi035087q

STRING, String Consortium 2017, http://string-db.org/cgi/network.pl?taskId=jwsLbfotgoq3

Sullivan, D. C, Huminiecki, L, Moore, J. W, Boyle, J. J, Poulsom, R, Creamer, D, Barker, J, Bicknell, R., 2003. EndoPDI, a novel protein-disulfide isomerase-like protein that is preferentially expressed in endothelial cells acts as a stress survival factor. J Biol Chem, 278(47):47079-88. DOI:10.1074/jbc.M308124200

Tokuhiro, K, Ikawa, M, Benham, A. M, Okabe, M., 2012. Protein disulfide isomerase homolog PDILT is required for quality control of sperm membrane protein ADAM3 and male fertility. Proceedings of the National Academy of Sciences 109(10): 3850-5. DOI: 10.1073/pnas.1117963109

Walker, W. H, Cheng, J., 2005. FSH and testosterone signalling in Sertoli cells. Reproduction 130:15-28. DOI: 10.1530/rep.1.00358

Wang, M, Kaufman, R. J., 2014. The impact of the endoplasmic reticulum protein-folding environment on cancer development. Nature Reviews Cancer, 14(9): 581-97. DOI: 10.1038/nrc3800

Williams, D. B., 2005. Beyond lectins: the calnexin/calreticulin chaperone system of the endoplasmic reticulum. Journal of Cell Science 119(4):615-23. DOI: 10.1242/jcs.02856

Xu, C., 2005. Endoplasmic reticulum stress: cell life and death decisions. Journal of Clinical Investigation 115(10) 2656-64. DOI: 10.1172/JCI26373

Yamaguchi, R, Yamagata, K, Ikawa, M, Moss, S. B, Okabe, M., 2006. Aberrant

distribution of ADAM3 in Sperm from both Angiotensin-converting enzyme (Ace)- and Calmegin (Clgn)- deficient mice. Biology of reproduction, 75(5): 760-766. DOI: 10.1095/biolreprod.106.052977

Yuan. R, Primakoff, P, Myles, D. G., 1997. A Role for the Disintegrin domain of Cyritestin, a Sperm surface protein belonging to the ADAM Family, in Mouse Sperm-Egg Plasma Membrane Adhesion and Fusion. The Journal of Cell Biology, 137(1):105-12f. DOI: 10.1083/jcb.137.1.105

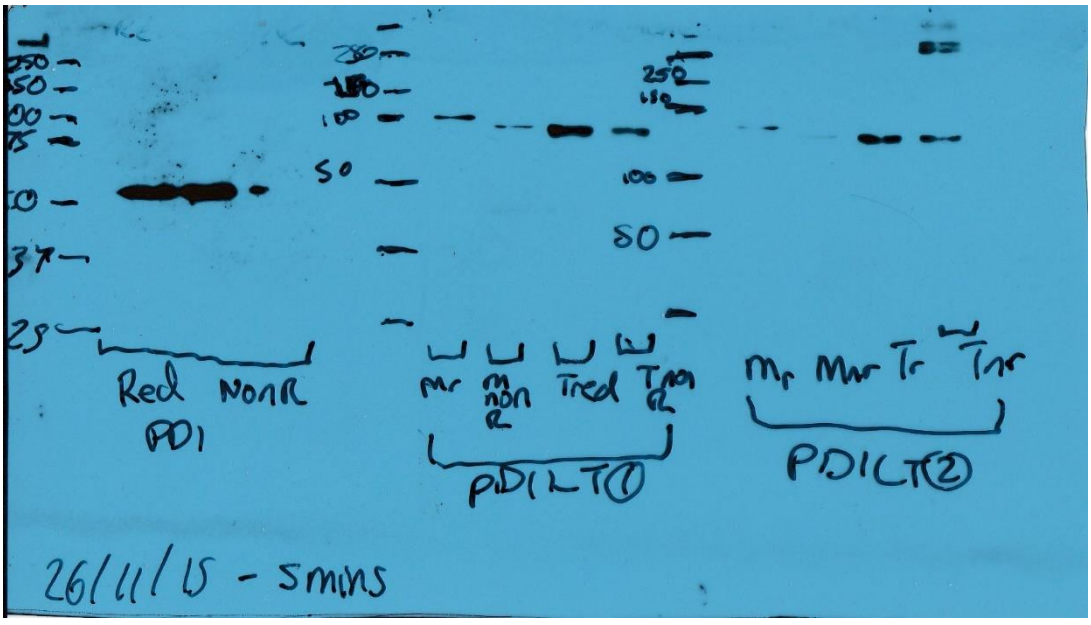
6. Appendix

Appendix 1: figure 14 – ERP44 full alignment

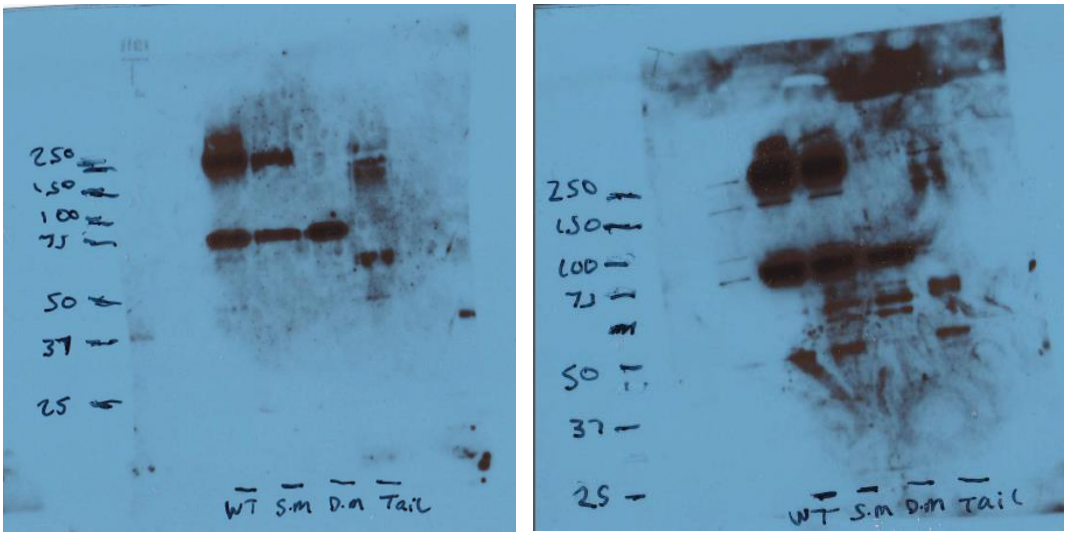
```
Q9BS26 ERP44_HUMAN1    MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEITSLDTENIDEILNNADVALVNFYADWCRF   60
H2QXK9 H2QXK9_PANTR1   MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEITSLDTENIDEILNNADVALVNFYADWCRF   60

Q9BS26 ERP44_HUMAN61   SQMLHPIFEEASDVIKEEFPNENQVVFARVDCDQHSDIAQRYRISKYPTLKLFRNGMMMK  120
H2QXK9 H2QXK9_PANTR61  SQMLHPIFEEASDVIKEEFPNENQVVFARVDCDQHSDIAQRYRISKYPTLKLFRNGMMMK  120

Q9BS26 ERP44_HUMAN121  REYRGQRSVKALADYIRQQKSDPIQEIRDLAEITTLDRSKRNIIGYFEQKDSDNYRVFER  180
H2QXK9 H2QXK9_PANTR121 REYRGQRSVKALADYIRQQKSDPIQEIRDLAEITTLDRSKRNIIGYFEQKDSDNYRVFER  180

Q9BS26 ERP44_HUMAN181  VANILHDDCAFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNWIQDK  240
H2QXK9 H2QXK9_PANTR181 VANILHDDCAFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNWIQDK  240

Q9BS26 ERP44_HUMAN241  CVPLVREITFENGEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHAD  300
H2QXK9 H2QXK9_PANTR241 CVPLVREITFENGEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHAD  300

Q9BS26 ERP44_HUMAN301  CDKFRHPLLHIQKTPADCPVIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDLHSGKLHREF  360
H2QXK9 H2QXK9_PANTR301 CDKFRHPLLHIQKTPADCPVIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDLHSGKLHREF  360

Q9BS26 ERP44_HUMAN361  HHGPDPTDTAPGEQAQDVASSPPESSFQKLAPSEYRYTLLRDRDEL               406
H2QXK9 H2QXK9_PANTR361 HHGPDPTDTAPGEQAQDVASSPPESSFQKLAPSEYRYTLLRDRDEL               406

Q9BS26 ERP44_HUMAN1    MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEITSLDTENIDEILNNADVALVNFYADWCRF   60
H9EUJ3 H9EUJ3_MACMU   MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEITSLDTENIDEILNNADVALVNFYADWCRF   60

Q9BS26 ERP44_HUMAN61   SQMLHPIFEEASDVIKEEFPNENQVVFARVDCDQHSDIAQRYRISKYPTLKLFRNGMMMK  120
H9EUJ3 H9EUJ3_MACMU   SQMLHPIFEEASDVIKEEFPNENQVVFARVDCDQHSDIAQRYRISKYPTLKLFRNGMMMK  120

Q9BS26 ERP44_HUMAN121  REYRGQRSVKALADYIRQQKSDPIQEIRDLAEITTLDRSKRNIIGYFEQKDSDNYRVFER  180
H9EUJ3 H9EUJ3_MACMU   REYRGQRSVKALADYIRQQKSDPIQEIRDLAEITTLDRSKRNIIGYFEQKDSDNYRVFER  180

Q9BS26 ERP44_HUMAN181  VANILHDDCAFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNWIQDK  240
H9EUJ3 H9EUJ3_MACMU   VANILHDDCAFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNWIQDK  240

Q9BS26 ERP44_HUMAN241  CVPLVREITFENGEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHAD  300
H9EUJ3 H9EUJ3_MACMU   CVPLVREITFENGEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHAD  300

Q9BS26 ERP44_HUMAN301  CDKFRHPLLHIQKTPADCPVIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDLHSGKLHREF  360
H2QXK9 H9EUJ3_MACMU    CDKFRHPLLHIQKTPADCPVIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDLHSGKLHREF  360

Q9BS26 ERP44_HUMAN361  HHGPDPTDTAPGEQAQDVASSPPESSFQKLAPSEYRYTLLRDRDEL               406
H2QXK9 H9EUJ3_MACMU    HHGPDPTDTAPGEQAQDVASSPPESSFQKLAPSEYRYTLLRDRDEL               406


Q9BS26 ERP44_HUMAN6    MHPAVFLSLPDLRCSLLLL-VTWVFTPVTTEITSLDTENIDEILNNADVALVNFYADWCRF  60
F7H0S1 F7H0S1_CALJA5   MHPAVFASLVSLSCNTLLSPVTWVFTPVTTEITSLDTENIDEILNNADVALVNFYADWCRF  60

Q9BS26 ERP44_HUMAN61   SQMLHPIFEEASDVIKEEFPNENQVVFARVDCDQHSDIAQRYRISKYPTLKLFRNGMMMK  120
H9EUJ3 F7H0S1_CALJA5  SQMLHPIFEEASDVIKEEFPNENQVVFARVDCDQHSDIAQRYRISKYPTLKLFRNGMMMK  120

Q9BS26 ERP44_HUMAN121  REYRGQRSVKALADYIRQQKSDPIQEIRDLAEITTLDRSKRNIIGYFEQKDSDNYRVFER  180
H9EUJ3 F7H0S1_CALJA5  REYRGQRSVKALADYIRQQKSDPIQEIRDLAEITTLDRSKRNIIGYFEQKDSDNYRVFER  180

Q9BS26 ERP44_HUMAN181  VANILHDDCAFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNWIQDK  240
H9EUJ3 F7H0S1_CALJA5  VANILHDDCAFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNWIQDK  240

Q9BS26 ERP44_HUMAN241  CVPLVREITFENGEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHAD  300
H9EUJ3 F7H0S1_CALJA5  CVPLVREITFENGEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHAD  300

Q9BS26 ERP44_HUMAN301  CDKFRHPLLHIQKTPADCPVIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDLHSGKLHREF  360
H2QXK9 F7H0S1_CALJA5  CDKFRHPLLHIQKTPADCPVIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDLHSGKLHREF  360

Q9BS26 ERP44_HUMAN361  HHGPDPTDTAPGEQAQDVASSPPESSFQKLAPSEYRYTLLRDRDEL               406
H2QXK9 F7H0S1_CALJA5  HHGPDPTDTAPGEQAQDVASSPPESSFQKLAPSEYRYTLLRDRDEL               406
```

```
Q9BS26 ERP44_HUMAN1   MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEITSLDTENIDEILNNADVALVNFYADWCRF   60
G5AZ30 G5AZ30_HETGA1 MNPAVFLSLPDLRCSLLLLVSWIFTPVTTEITSLDTENIDEILNNADVALVNFYADWCRF   60

Q9BS26 ERP44_HUMAN61  SQMLHPIFEEASDVIKEEFPNENQVVFARVDCDQHSDIAQRYRISKYPTLKLFRNGMMMK  120
G5AZ30 G5AZ30_HETGA1 SQMLHPIFEEASDVIKEEYPNQNQVVFARVDCDQHSDIAQRYRISKYPTLKLFRNGMMMK  120
Q9BS26 ERP44_HUMAN121 REYRGQRSVKALADYIRQQKSDPIQEIRDLAEITTLDRSKRNIIGYFEQKDSDNYRVFER  180
G5AZ30 G5AZ30_HETGA1 REYRGQRSVKALADYIRQQKSDPVQEMQNLEEITTLDRSKRNIIGFFEQKDSADYKVFER  180

Q9BS26 ERP44_HUMAN181 VANILHDDCAFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNWIQDK  240
G5AZ30 G5AZ30_HETGA1 VANILHDDCAFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNWIQDK  240

Q9BS26 ERP44_HUMAN241 CVPLVREITFENGEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHAD  300
G5AZ30 G5AZ30_HETGA1 CVPLVREITFENGEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHAD  300

Q9BS26 ERP44_HUMAN301 CDKFRHPLLHIQKTPADCPVIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDLHSGKLHREF  360
G5AZ30 G5AZ30_HETGA1 CDKFRHPLLHIQKTPADCPVIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDLHSGKLHREF  360

Q9BS26 ERP44_HUMAN361 HHGPDPTDTAPGEQAQDVASSPPESSFQKLAPSEYRYTLLRDRDEL               406
G5AZ30 G5AZ30_HETGA1 HHGPDPTDTAPGEQAQDVASSPPESSFQKLAPSEYRYTLLRDRDEL               406
```

Appendix 2: figure 20 – Western blot of Mock and PDILT lysates under reducing and non-reducing conditions with lighter exposure (for 5 minutes)



Appendix 3: Western blot of the PDILT WT and mutant lysates, under non-reducing conditions. Blotted for α-PDILT. Other exposures of figure 26.



Appendix 4: Sequence alignment of TMX proteins

```
SP|Q96JJ7|TMX3_HUMAN ------------------------------------------------------------
SP|Q9H3N1|TMX1_HUMAN ------------------------------------------------------------
SP|Q9Y320|TMX2_HUMAN MAVLAPLIALVYSVPRLSRWLAQPYYLLSALLSAAFLLVRKLPPLCHGLPTQREDGNPCD 60
SP|Q9H1E5|TMX4_HUMAN ------------------------------------------------------------


SP|Q96JJ7|TMX3_HUMAN ------------------------------------------------MAAWKSWTALRLC 13
SP|Q9H3N1|TMX1_HUMAN ---------------------------------MAPSGSLAVPLAVLVLLLWGAPW----- 23
SP|Q9Y320|TMX2_HUMAN FDWREVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLC 120
SP|Q9H1E5|TMX4_HUMAN ---------------------------------MAGGRCGPQLTALLAAWIAAVAATAG 26
                                                                :    .


SP|Q96JJ7|TMX3_HUMAN A-------TVVVLDMVVCKGFVEDLDESFKENRNDDIWLVDFYAPWCGHCKKLEPIWNEV 66
SP|Q9H3N1|TMX1_HUMAN --------------THGRRSNVRVITDENWRELLEGDWMIEFYAPWCPACQNLQPEWESF 69
SP|Q9Y320|TMX2_HUMAN IVFLMTCKPPLYMGPEYIKYFNDKTIDEELERDKRVTWIVEFFANWSNDCQSFAPIYADL 180
SP|Q9H1E5|TMX4_HUMAN P-------EEA--ALPPEQSRVQPMTASNWTLVMEGEWMLKFYAPWCPSCQQTDSEWEAF 77
                                  :        .         *::.*:* *.  *:.    :  .


SP|Q96JJ7|TMX3_HUMAN GLEMKSIGSPVKVGKMDATSYSSIASEFGVRGYPTIKLLKGDLAY-----NYRGPRTKDD 121
SP|Q9H3N1|TMX1_HUMAN AEWGE--DLEVNIAKVDVTEQPGLSGRFIITALPTIYHCKDGEFR-----RYQGPRTKKD 122
SP|Q9Y320|TMX2_HUMAN SLKYN--CTGLNFGKVDVGRYTDVSTRYKVSTSPLTKQLPTLILFQGGKEAMRRPQIDKK 238
SP|Q9H1E5|TMX4_HUMAN AKNGE--ILQISVGKVDVIQEPGLSGRFFVTTLPAFFHAKDGIFR-----RYRGPGIFED 130
                    .   :     :...*:*.    .:: .: :   *                : *   ..
```

```
SP|Q96JJ7|TMX3_HUMAN RERFQNYLAMDGFLLYELGDTGKLVALAVIDEKNTSVEHTRLKSIIQEVARDYRDLF--- 281
SP|Q9H3N1|TMX1_HUMAN KK-L---LSESA---QPLK--------KVEEEQEAD----------------EEDVSEEE 250
SP|Q9Y320|TMX2_HUMAN KKDK------------------------------------------------------- 296
SP|Q9H1E5|TMX4_HUMAN RRSE---EAHRA---EQLQ--------DAEEEKDDSNEEENKDSLVDDE-EEKEDLGDED 274
                     :.

SP|Q96JJ7|TMX3_HUMAN -------HRDFQFGH----MDGNDYINTLLMDELTV----PTVVVLNTSNQQYFLLDRQI 326
SP|Q9H3N1|TMX1_HUMAN AESKEGTNKDFPQNA----IRQRSLGP-SLATDKS------------------------ 280
SP|Q9Y320|TMX2_HUMAN ------------------------------------------------------------
SP|Q9H1E5|TMX4_HUMAN EAEEEEEEDNLAAGVDEERSEANDQGP-PGEDGVTREEVEPEEAEEGISEQP-CPAD--- 329


SP|Q96JJ7|TMX3_HUMAN KNVEDMVQFINNILDGTVEAQGGDSILQRLKRIVFDAKSTIVSIFKSSPLMGCFLFGLPL 386
SP|Q9H3N1|TMX1_HUMAN ------------------------------------------------------------
SP|Q9Y320|TMX2_HUMAN ------------------------------------------------------------
SP|Q9H1E5|TMX4_HUMAN ----------TEVVE--------DSLRQRKSQ---HADKGL------------------- 349


SP|Q96JJ7|TMX3_HUMAN GVISIMCYGIYTADTDGGYIEERYEVSKSENENQEQIEESKEQQEPSSGGSVVPTVQEPK 446
SP|Q9H3N1|TMX1_HUMAN ------------------------------------------------------------
SP|Q9Y320|TMX2_HUMAN ------------------------------------------------------------
SP|Q9H1E5|TMX4_HUMAN ------------------------------------------------------------

SP|Q96JJ7|TMX3_HUMAN DVLEKKKD 454
SP|Q9H3N1|TMX1_HUMAN --------
SP|Q9Y320|TMX2_HUMAN --------
SP|Q9H1E5|TMX4_HUMAN --------
```